```
DR   PRINTS; PR01541; GONADOLIBRNI.
DR   PROSITE; PS00473; GNRH; 1.
KW   Amidation; Hormone; Signal.
FT   SIGNAL        1     26       POTENTIAL.
FT   CHAIN        27     36       SEABREAM-TYPE GONADOTROPIN-RELEASING
FT                                HORMONE.
FT   CHAIN        40     96       GONADOTROPIN-RELEASING HORMONE ASSOCIATED
FT                                PEPTIDE.
SQ   SEQUENCE   96 AA;  10560 MW;  FA3202565EA00DCC CRC64;

  Query Match             32.1%;  Score 52;  DB 13;  Length 96;
  Best Local Similarity   88.9%;  Pred. No. 7.6;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 HWSYGLRPG 10
              |||||| ||
Db         28 HWSYGLSPG 36


RESULT 15
Q805A5
ID   Q805A5      PRELIMINARY;      PRT;    98 AA.
AC   Q805A5;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Seabream-type gonadotropin-releasing hormone precursor.
OS   Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidei;
OC   Cichlidae; Oreochromis.
OX   NCBI_TaxID=8128;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RA   Parhar I.S., Ogawa S., Sakuma Y.;
RT   "Molecular cloning of tilapia (Oreochromis niloticus) GnRH cDNA.";
RL   Submitted (FEB-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB101665; BAC56849.1; -.
DR   GO; GO:0005576; C:extracellular; IEA.
DR   GO; GO:0005183; F:luteinizing hormone-releasing factor activity; IEA.
DR   GO; GO:0007275; P:development; IEA.
DR   InterPro; IPR002012; GnRH.
DR   InterPro; IPR004079; GonadoliberinI.
DR   Pfam; PF00446; GnRH; 1.
DR   PRINTS; PR01541; GONADOLIBRNI.
DR   PROSITE; PS00473; GNRH; 1.
KW   Signal.
FT   SIGNAL        1     22       POTENTIAL.
FT   CHAIN        23     32       SEABREAM-TYPE GONADOTROPIN-RELEASING
FT                                HORMONE.
FT   CHAIN        36     98       SEABREAM-TYPE GNRH ASSOCIATED PEPTIDE.
SQ   SEQUENCE   98 AA;  10741 MW;  2417F7F37DBC97E3 CRC64;

  Query Match             32.1%;  Score 52;  DB 13;  Length 98;
  Best Local Similarity   88.9%;  Pred. No. 7.8;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 HWSYGLRPG 10
              |||||| ||
Db         24 HWSYGLSPG 32


Search completed: March 10, 2004, 09:25:38
Job time : 32.6381 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:48 ; Search time 46.6809 Seconds
(without alignments)
187.635 Million cell updates/sec

Title: US-09-848-834A-15
Perfect score: 162
Sequence: 1 XHWSYGLRPGSSGPSLQYIKANSKFIGITEL 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	161	99.4	31	5	AAU11426	Aau11426 Synthetic
2	161	99.4	46	5	AAU11430	Aau11430 Synthetic
3	89	54.9	34	5	AAU11424	Aau11424 Synthetic
4	89	54.9	36	5	AAU11427	Aau11427 Synthetic
5	89	54.9	47	5	AAU11428	Aau11428 Synthetic
6	89	54.9	51	5	AAU11431	Aau11431 Synthetic
7	87	53.7	37	5	AAU11425	Aau11425 Synthetic
8	87	53.7	50	5	AAU11429	Aau11429 Synthetic
9	85	52.5	109	4	AAB20147	Aab20147 Growth di
10	84	51.9	24	5	ABP51515	Abp51515 HBV antig
11	83	51.2	216	3	AAY92665	Aay92665 MUC-1 ana
12	82.5	50.9	158	2	AAW81331	Aaw81331 TNF2-7, a
13	82.5	50.9	158	5	ABB07277	Abb07277 Human TNF
14	81	50.0	116	3	AAB45502	Aab45502 Modified
15	81	50.0	116	3	AAB45526	Aab45526 Modified
16	80.5	49.7	158	2	AAW81327	Aaw81327 TNF2-1, a
17	80.5	49.7	158	5	ABB07280	Abb07280 Human TNF
18	79	48.8	118	3	AAB45491	Aab45491 Modified
19	79	48.8	118	3	AAB45518	Aab45518 Modified
20	79	48.8	287	6	AAO30460	Aao30460 hIL5.37 v
21	78	48.1	72	4	AAB46190	Aab46190 Tetanus t
22	78	48.1	136	4	AAB49089	Aab49089 Amyloid b
23	78	48.1	182	3	AAY84424	Aay84424 An osteop
24	77	47.5	173	3	AAY84425	Aay84425 DNA encod
25	77	47.5	194	6	AAO30488	Aao30488 Human TNF
26	76	46.9	16	5	AAU11413	Aau11413 Tetanus t
27	76	46.9	17	6	ADA09238	Ada09238 Tetanus t
28	76	46.9	19	3	AAY99055	Aay99055 HLA class
29	76	46.9	37	2	AAR65389	Aar65389 Universal
30	76	46.9	37	2	AAR65383	Aar65383 Universal
31	76	46.9	47	2	AAR62723	Aar62723 LHRH-cont
32	76	46.9	124	3	AAB45492	Aab45492 Modified
33	76	46.9	124	3	AAB45517	Aab45517 Modified
34	76	46.9	160	4	AAB20153	Aab20153 Growth di
35	76	46.9	194	6	AAO30489	Aao30489 Human TNF
36	76	46.9	285	6	AAO30457	Aao30457 hIL5-P30-
37	76	46.9	573	1	AAP70345	Aap70345 Portion o
38	76	46.9	1315	4	AAB61169	Aab61169 Clostridi
39	75	46.3	25	3	AAY92650	Aay92650 PSMpep007
40	75	46.3	693	3	AAY92647	Aay92647 Mutant hu
41	75	46.3	750	3	AAY92637	Aay92637 Mutant hu
42	75	46.3	750	3	AAY92628	Aay92628 Mutant hu
43	75	46.3	750	3	AAY92644	Aay92644 Mutant hu
44	74	45.7	15	2	AAR06310	Aar06310 Tetanus t
45	74	45.7	15	2	ADB87354	Adb87354 Cytotoxic

ALIGNMENTS

RESULT 1
AAU11426
ID AAU11426 standard; peptide; 31 AA.
XX
AC AAU11426;
XX
DT 12-MAR-2002 (first entry)
XX
DE Synthetic immunogen peptide 7.
XX
KW Gonadotrophin releasing hormone; GnRH; synthetic immunogen;
KW luteinising hormone releasing hormone; LHRH; contraceptive;
KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;
KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.
XX
OS Clostridium tetani.
OS Mammalia.
OS Synthetic.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT Peptide 1. .10
FT /note= "Gonadotrophin releasing hormone epitope"
FT Misc-difference 1
FT /label= OTHER
FT /note= "Other= Pyro-glutamic acid or 5-oxo proline"
FT Peptide 11. .16
FT /note= "Spacer peptide"
FT Peptide 17. .31
FT /note= "Tetanus toxoid sequence (830-844 aa)"
XX
PN WO200185763-A2.
XX
PD 15-NOV-2001.
XX
PF 04-MAY-2001; 2001WO-US014363.
XX
PR 05-MAY-2000; 2000US-0202328P.
XX
PA (APHT-) APHTON CORP.
XX
PI Grimes S, Michaeli D, Stevens VC;
XX
DR WPI; 2002-049440/06.
XX
PT Novel synthetic immunogen for inducing immune response against
PT gonadotropin releasing hormone, comprises fusion peptide having

```
PT   promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
PT   or its analog.
XX
PS   Claim 11; Page 10; 43pp; English.
XX
CC   The invention relates to a synthetic immunogen for inducing specific
CC   antibodies against gonadotropin releasing hormone (GnRH also known as
CC   luteinising hormone releasing hormone, LHRH) comprising a fusion peptide
CC   which comprises a promiscuous helper T-cell peptide epitope and
CC   immunomimic peptide epitope or its analogue. The synthetic immunogen is
CC   useful inducing an immune response against GnRH in an animal subject, and
CC   as such is useful as a contraceptive and in the treatment of diseases
CC   such as cancer (of the breast, uterus and other gynaecological cancer),
CC   endometriosis, uterine fibroids, benign prostatic hypertrophy and
CC   prostate cancer. The immunogen is effective in eliciting high and
CC   specific anti-GnRH antibody titres. The present sequence is a synthetic
CC   immunogen of the invention
XX
SQ   Sequence 31 AA;

  Query Match             99.4%;  Score 161;  DB 5;  Length 31;
  Best Local Similarity   100.0%;  Pred. No. 3e-16;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 HWSYGLRPGSSGPSLQYIKANSKFIGITEL 31
              ||||||||||||||||||||||||||||||
Db          2 HWSYGLRPGSSGPSLQYIKANSKFIGITEL 31


RESULT 2
AAU11430
ID   AAU11430 standard; peptide; 46 AA.
XX
AC   AAU11430;
XX
DT   12-MAR-2002  (first entry)
XX
DE   Synthetic immunogen peptide 11.
XX
KW   Gonadotrophin releasing hormone; GnRH; synthetic immunogen;
KW   luteinising hormone releasing hormone; LHRH; contraceptive;
KW   promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
KW   breast cancer; uterine cancer; gynaecological cancer; endometriosis;
KW   uterine fibroid; benign prostatic hypertrophy; prostate cancer.
XX
OS   Clostridium tetani.
OS   Mammalia.
OS   Synthetic.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .10
FT                   /note= "Gonadotrophin releasing hormone epitope (1. .10
FT                   aa)"
FT   Misc-difference 1
FT                   /label= OTHER
FT                   /note= "Other= Pyro-glutamic acid or 5-oxo proline"
FT   Peptide         11. .16
FT                   /note= "Spacer peptide"
FT   Peptide         17. .31
FT                   /note= "Tetanus toxoid (830-844 aa)"
FT   Peptide         32. .37
FT                   /note= "Spacer peptide"
FT   Peptide         38. .46
FT                   /note= "Gonadotrophin releasing hormone epitope (2-10
FT                   aa)"
FT   Modified-site   46
FT                   /note= "Amidated glycine or glycinamide"
XX
PN   WO200185763-A2.
XX
PD   15-NOV-2001.
XX
PF   04-MAY-2001; 2001WO-US014363.
XX
PR   05-MAY-2000; 2000US-0202328P.
XX
PA   (APHT-) APHTON CORP.
XX
PI   Grimes S,  Michaeli D,  Stevens VC;
XX
DR   WPI; 2002-049440/06.
XX
PT   Novel synthetic immunogen for inducing immune response against
PT   gonadotropin releasing hormone, comprises fusion peptide having
PT   promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
PT   or its analog.
XX
PS   Claim 11; Page 12; 43pp; English.
XX
CC   The invention relates to a synthetic immunogen for inducing specific
CC   antibodies against gonadotropin releasing hormone (GnRH also known as
CC   luteinising hormone releasing hormone, LHRH) comprising a fusion peptide
CC   which comprises a promiscuous helper T-cell peptide epitope and
CC   immunomimic peptide epitope or its analogue. The synthetic immunogen is
CC   useful inducing an immune response against GnRH in an animal subject, and
CC   as such is useful as a contraceptive and in the treatment of diseases
CC   such as cancer (of the breast, uterus and other gynaecological cancer),
CC   endometriosis, uterine fibroids, benign prostatic hypertrophy and
CC   prostate cancer. The immunogen is effective in eliciting high and
CC   specific anti-GnRH antibody titres. The present sequence is a synthetic
CC   immunogen of the invention
XX
SQ   Sequence 46 AA;

  Query Match             99.4%;  Score 161;  DB 5;  Length 46;
  Best Local Similarity   100.0%;  Pred. No. 4.8e-16;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 HWSYGLRPGSSGPSLQYIKANSKFIGITEL 31
              ||||||||||||||||||||||||||||||
Db          2 HWSYGLRPGSSGPSLQYIKANSKFIGITEL 31


RESULT 3
AAU11424
ID   AAU11424 standard; peptide; 34 AA.
XX
AC   AAU11424;
XX
DT   12-MAR-2002  (first entry)
XX
DE   Synthetic immunogen peptide 5.
XX
KW   Gonadotrophin releasing hormone; GnRH; synthetic immunogen;
KW   luteinising hormone releasing hormone; LHRH; contraceptive;
KW   promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
KW   breast cancer; uterine cancer; gynaecological cancer; endometriosis;
KW   uterine fibroid; benign prostatic hypertrophy; prostate cancer.
XX
OS   Measles virus.
OS   Mammalia.
OS   Synthetic.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .10
FT                   /note= "Gonadotrophin releasing hormone epitope"
FT   Misc-difference 1
FT                   /label= OTHER
FT                   /note= "Other= Pyro-glutamic acid or 5-oxo proline"
FT   Peptide         11. .16
FT                   /note= "Spacer peptide"
FT   Peptide         17. .34
FT                   /note= "Measles virus fusion protein F epitope"
XX
```

```
PN   WO200185763-A2.
XX
PD   15-NOV-2001.
XX
PF   04-MAY-2001; 2001WO-US014363.
XX
PR   05-MAY-2000; 2000US-0202328P.
XX
PA   (APHT-) APHTON CORP.
XX
PI   Grimes S,  Michaeli D,  Stevens VC;
XX
DR   WPI; 2002-049440/06.
XX
PT   Novel synthetic immunogen for inducing immune response against
PT   gonadotropin releasing hormone, comprises fusion peptide having
PT   promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
PT   or its analog.
XX
PS   Claim 11; Page 9; 43pp; English.
XX
CC   The invention relates to a synthetic immunogen for inducing specific
CC   antibodies against gonadotropin releasing hormone (GnRH also known as
CC   luteinising hormone releasing hormone, LHRH) comprising a fusion peptide
CC   which comprises a promiscuous helper T-cell peptide epitope and
CC   immunomimic peptide epitope or its analogue. The synthetic immunogen is
CC   useful inducing an immune response against GnRH in an animal subject, and
CC   as such is useful as a contraceptive and in the treatment of diseases
CC   such as cancer (of the breast, uterus and other gynaecological cancer),
CC   endometriosis, uterine fibroids, benign prostatic hypertrophy and
CC   prostate cancer. The immunogen is effective in eliciting high and
CC   specific anti-GnRH antibody titres. The present sequence is a synthetic
CC   immunogen of the invention
XX
SQ   Sequence 34 AA;

  Query Match             54.9%;  Score 89;  DB 5;  Length 34;
  Best Local Similarity   83.3%;  Pred. No. 1.2e-05;
  Matches   15;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 HWSYGLRPGSSGPSLQYI 19
              |||||||||||||||: :
Db          2 HWSYGLRPGSSGPSLKLL 19


RESULT 4
AAU11427
ID   AAU11427 standard; peptide; 36 AA.
XX
AC   AAU11427;
XX
DT   12-MAR-2002  (first entry)
XX
DE   Synthetic immunogen peptide 8.
XX
KW   Gonadotrophin releasing hormone; GnRH; synthetic immunogen;
KW   luteinising hormone releasing hormone; LHRH; contraceptive;
KW   promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
KW   breast cancer; uterine cancer; gynaecological cancer; endometriosis;
KW   uterine fibroid; benign prostatic hypertrophy; prostate cancer.
XX
OS   Plasmodium falciparum.
OS   Mammalia.
OS   Synthetic.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .10
FT                   /note= "Gonadotrophin releasing hormone epitope"
FT   Misc-difference 1
FT                   /label= OTHER
FT                   /note= "Other= Pyro-glutamic acid or 5-oxo proline"
FT   Peptide         11. .16
FT                   /note= "Spacer peptide"
FT   Peptide         17. .36
FT                   /note= "Malaria CSP protein (378-398 aa)"
XX
PN   WO200185763-A2.
XX
PD   15-NOV-2001.
XX
PF   04-MAY-2001; 2001WO-US014363.
XX
PR   05-MAY-2000; 2000US-0202328P.
XX
PA   (APHT-) APHTON CORP.
XX
PI   Grimes S,  Michaeli D,  Stevens VC;
XX
DR   WPI; 2002-049440/06.
XX
PT   Novel synthetic immunogen for inducing immune response against
PT   gonadotropin releasing hormone, comprises fusion peptide having
PT   promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
PT   or its analog.
XX
PS   Claim 11; Page 10; 43pp; English.
XX
CC   The invention relates to a synthetic immunogen for inducing specific
CC   antibodies against gonadotropin releasing hormone (GnRH also known as
CC   luteinising hormone releasing hormone, LHRH) comprising a fusion peptide
CC   which comprises a promiscuous helper T-cell peptide epitope and
CC   immunomimic peptide epitope or its analogue. The synthetic immunogen is
CC   useful inducing an immune response against GnRH in an animal subject, and
CC   as such is useful as a contraceptive and in the treatment of diseases
CC   such as cancer (of the breast, uterus and other gynaecological cancer),
CC   endometriosis, uterine fibroids, benign prostatic hypertrophy and
CC   prostate cancer. The immunogen is effective in eliciting high and
CC   specific anti-GnRH antibody titres. The present sequence is a synthetic
CC   immunogen of the invention
XX
SQ   Sequence 36 AA;

  Query Match             54.9%;  Score 89;  DB 5;  Length 36;
  Best Local Similarity   63.3%;  Pred. No. 1.3e-05;
  Matches   19;  Conservative    2;  Mismatches    3;  Indels    6;  Gaps    1;

Qy          2 HWSYGLRPGSSGPSL------QYIKANSKF 25
              |||||||||||||||      :  ||:| |
Db          2 HWSYGLRPGSSGPSLDEKKIAKMEKASSVF 31


RESULT 5
AAU11428
ID   AAU11428 standard; peptide; 47 AA.
XX
AC   AAU11428;
XX
DT   12-MAR-2002  (first entry)
XX
DE   Synthetic immunogen peptide 9.
XX
KW   Gonadotrophin releasing hormone; GnRH; synthetic immunogen;
KW   luteinising hormone releasing hormone; LHRH; contraceptive;
KW   promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
KW   breast cancer; uterine cancer; gynaecological cancer; endometriosis;
KW   uterine fibroid; benign prostatic hypertrophy; prostate cancer.
XX
OS   Plasmodium falciparum.
OS   Mammalia.
OS   Synthetic.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .10
FT                   /note= "Gonadotrophin releasing hormone epitope (1. .10
```

```
FT                  aa)"
FT   Misc-difference 1
FT                  /label= OTHER
FT                  /note= "Other= Pyro-glutamic acid or 5-oxo proline"
FT   Peptide        11. .16
FT                  /note= "Spacer peptide"
FT   Peptide        17. .34
FT                  /note= "Malaria CSP protein (288-302 aa)"
FT   Peptide        35. .38
FT                  /note= "Spacer peptide"
FT   Peptide        39. .47
FT                  /note= "Gonadotrophin releasing hormone epitope (2-10
FT                  aa)"
FT   Modified-site  47
FT                  /note= "Amidated glycine or glycinamide"
XX
PN   WO200185763-A2.
XX
PD   15-NOV-2001.
XX
PF   04-MAY-2001; 2001WO-US014363.
XX
PR   05-MAY-2000; 2000US-0202328P.
XX
PA   (APHT-) APHTON CORP.
XX
PI   Grimes S,  Michaeli D,  Stevens VC;
XX
DR   WPI; 2002-049440/06.
XX
PT   Novel synthetic immunogen for inducing immune response against
PT   gonadotropin releasing hormone, comprises fusion peptide having
PT   promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
PT   or its analog.
XX
PS   Claim 11; Page 11; 43pp; English.
XX
CC   The invention relates to a synthetic immunogen for inducing specific
CC   antibodies against gonadotropin releasing hormone (GnRH also known as
CC   luteinising hormone releasing hormone, LHRH) comprising a fusion peptide
CC   which comprises a promiscuous helper T-cell peptide epitope and
CC   immunomimic peptide epitope or its analogue. The synthetic immunogen is
CC   useful inducing an immune response against GnRH in an animal subject, and
CC   as such is useful as a contraceptive and in the treatment of diseases
CC   such as cancer (of the breast, uterus and other gynaecological cancer),
CC   endometriosis, uterine fibroids, benign prostatic hypertrophy and
CC   prostate cancer. The immunogen is effective in eliciting high and
CC   specific anti-GnRH antibody titres. The present sequence is a synthetic
CC   immunogen of the invention
XX
SQ   Sequence 47 AA;

  Query Match             54.9%;  Score 89;  DB 5;  Length 47;
  Best Local Similarity   83.3%;  Pred. No. 1.8e-05;
  Matches   15;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 HWSYGLRPGSSGPSLQYI 19
              |||||||||||||||: :
Db          2 HWSYGLRPGSSGPSLKLL 19


RESULT 6
AAU11431
ID   AAU11431 standard; peptide; 51 AA.
XX
AC   AAU11431;
XX
DT   12-MAR-2002  (first entry)
XX
DE   Synthetic immunogen peptide 12.
XX
KW   Gonadotrophin releasing hormone; GnRH; synthetic immunogen;
KW   luteinising hormone releasing hormone; LHRH; contraceptive;
KW   promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
KW   breast cancer; uterine cancer; gynaecological cancer; endometriosis;
KW   uterine fibroid; benign prostatic hypertrophy; prostate cancer.
XX
OS   Plasmodium falciparum.
OS   Mammalia.
OS   Synthetic.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .10
FT                   /note= "Gonadotrophin releasing hormone epitope (1. .10
FT                   aa)"
FT   Misc-difference 1
FT                   /label= OTHER
FT                   /note= "Other= Pyro-glutamic acid or 5-oxo proline"
FT   Peptide         11. .16
FT                   /note= "Spacer peptide"
FT   Peptide         17. .36
FT                   /note= "Malaria CSP protein (378-398 aa)"
FT   Peptide         37. .42
FT                   /note= "Spacer peptide"
FT   Peptide         43. .51
FT                   /note= "Gonadotrophin releasing hormone epitope (2-10
FT                   aa)"
FT   Modified-site   51
FT                   /note= "Amidated glycine or glycinamide"
XX
PN   WO200185763-A2.
XX
PD   15-NOV-2001.
XX
PF   04-MAY-2001; 2001WO-US014363.
XX
PR   05-MAY-2000; 2000US-0202328P.
XX
PA   (APHT-) APHTON CORP.
XX
PI   Grimes S,  Michaeli D,  Stevens VC;
XX
DR   WPI; 2002-049440/06.
XX
PT   Novel synthetic immunogen for inducing immune response against
PT   gonadotropin releasing hormone, comprises fusion peptide having
PT   promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
PT   or its analog.
XX
PS   Claim 11; Page 12-13; 43pp; English.
XX
CC   The invention relates to a synthetic immunogen for inducing specific
CC   antibodies against gonadotropin releasing hormone (GnRH also known as
CC   luteinising hormone releasing hormone, LHRH) comprising a fusion peptide
CC   which comprises a promiscuous helper T-cell peptide epitope and
CC   immunomimic peptide epitope or its analogue. The synthetic immunogen is
CC   useful inducing an immune response against GnRH in an animal subject, and
CC   as such is useful as a contraceptive and in the treatment of diseases
CC   such as cancer (of the breast, uterus and other gynaecological cancer),
CC   endometriosis, uterine fibroids, benign prostatic hypertrophy and
CC   prostate cancer. The immunogen is effective in eliciting high and
CC   specific anti-GnRH antibody titres. The present sequence is a synthetic
CC   immunogen of the invention
XX
SQ   Sequence 51 AA;

  Query Match             54.9%;  Score 89;  DB 5;  Length 51;
  Best Local Similarity   63.3%;  Pred. No. 1.9e-05;
  Matches   19;  Conservative    2;  Mismatches    3;  Indels    6;  Gaps    1;

Qy          2 HWSYGLRPGSSGPSL------QYIKANSKF 25
              |||||||||||||||      :  ||:| |
Db          2 HWSYGLRPGSSGPSLDEKKIAKMEKASSVF 31
```

```
RESULT 7
AAU11425
ID   AAU11425 standard; peptide; 37 AA.
XX
AC   AAU11425;
XX
DT   12-MAR-2002  (first entry)
XX
DE   Synthetic immunogen peptide 6.
XX
KW   Gonadotrophin releasing hormone; GnRH; synthetic immunogen;
KW   luteinising hormone releasing hormone; LHRH; contraceptive;
KW   promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
KW   breast cancer; uterine cancer; gynaecological cancer; endometriosis;
KW   uterine fibroid; benign prostatic hypertrophy; prostate cancer.
XX
OS   Clostridium tetani.
OS   Mammalia.
OS   Synthetic.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .10
FT                   /note= "Gonadotrophin releasing hormone epitope"
FT   Misc-difference 1
FT                   /label= OTHER
FT                   /note= "Other= Pyro-glutamic acid or 5-oxo proline"
FT   Peptide         11. .16
FT                   /note= "Spacer peptide"
FT   Peptide         17. .37
FT                   /note= "Tetanus toxoid sequence (947-967 aa)"
XX
PN   WO200185763-A2.
XX
PD   15-NOV-2001.
XX
PF   04-MAY-2001; 2001WO-US014363.
XX
PR   05-MAY-2000; 2000US-0202328P.
XX
PA   (APHT-) APHTON CORP.
XX
PI   Grimes S,  Michaeli D,  Stevens VC;
XX
DR   WPI; 2002-049440/06.
XX
PT   Novel synthetic immunogen for inducing immune response against
PT   gonadotropin releasing hormone, comprises fusion peptide having
PT   promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
PT   or its analog.
XX
PS   Claim 11; Page 9; 43pp; English.
XX
CC   The invention relates to a synthetic immunogen for inducing specific
CC   antibodies against gonadotropin releasing hormone (GnRH also known as
CC   luteinising hormone releasing hormone, LHRH) comprising a fusion peptide
CC   which comprises a promiscuous helper T-cell peptide epitope and
CC   immunomimic peptide epitope or its analogue. The synthetic immunogen is
CC   useful inducing an immune response against GnRH in an animal subject, and
CC   as such is useful as a contraceptive and in the treatment of diseases
CC   such as cancer (of the breast, uterus and other gynaecological cancer),
CC   endometriosis, uterine fibroids, benign prostatic hypertrophy and
CC   prostate cancer. The immunogen is effective in eliciting high and
CC   specific anti-GnRH antibody titres. The present sequence is a synthetic
CC   immunogen of the invention
XX
SQ   Sequence 37 AA;

  Query Match             53.7%;  Score 87;  DB 5;  Length 37;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-05;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 HWSYGLRPGSSGPSL 16
              |||||||||||||||
Db          2 HWSYGLRPGSSGPSL 16


RESULT 8
AAU11429
ID   AAU11429 standard; peptide; 50 AA.
XX
AC   AAU11429;
XX
DT   12-MAR-2002  (first entry)
XX
DE   Synthetic immunogen peptide 10.
XX
KW   Gonadotrophin releasing hormone; GnRH; synthetic immunogen;
KW   luteinising hormone releasing hormone; LHRH; contraceptive;
KW   promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
KW   breast cancer; uterine cancer; gynaecological cancer; endometriosis;
KW   uterine fibroid; benign prostatic hypertrophy; prostate cancer.
XX
OS   Clostridium tetani.
OS   Mammalia.
OS   Synthetic.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .10
FT                   /note= "Gonadotrophin releasing hormone epitope (1. .10
FT                   aa)"
FT   Misc-difference 1
FT                   /label= OTHER
FT                   /note= "Other= Pyro-glutamic acid or 5-oxo proline"
FT   Peptide         11. .16
FT                   /note= "Spacer peptide"
FT   Peptide         17. .37
FT                   /note= "Tetanus toxoid (947-967 aa)"
FT   Peptide         38. .41
FT                   /note= "Spacer peptide"
FT   Peptide         42. .50
FT                   /note= "Gonadotrophin releasing hormone epitope (2-10
FT                   aa)"
FT   Modified-site   50
FT                   /note= "Amidated glycine or glycinamide"
XX
PN   WO200185763-A2.
XX
PD   15-NOV-2001.
XX
PF   04-MAY-2001; 2001WO-US014363.
XX
PR   05-MAY-2000; 2000US-0202328P.
XX
PA   (APHT-) APHTON CORP.
XX
PI   Grimes S,  Michaeli D,  Stevens VC;
XX
DR   WPI; 2002-049440/06.
XX
PT   Novel synthetic immunogen for inducing immune response against
PT   gonadotropin releasing hormone, comprises fusion peptide having
PT   promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
PT   or its analog.
XX
PS   Claim 11; Page 11; 43pp; English.
XX
CC   The invention relates to a synthetic immunogen for inducing specific
CC   antibodies against gonadotropin releasing hormone (GnRH also known as
CC   luteinising hormone releasing hormone, LHRH) comprising a fusion peptide
CC   which comprises a promiscuous helper T-cell peptide epitope and
CC   immunomimic peptide epitope or its analogue. The synthetic immunogen is
CC   useful inducing an immune response against GnRH in an animal subject, and
CC   as such is useful as a contraceptive and in the treatment of diseases
CC   such as cancer (of the breast, uterus and other gynaecological cancer),
```

```
CC   endometriosis, uterine fibroids, benign prostatic hypertrophy and
CC   prostate cancer. The immunogen is effective in eliciting high and
CC   specific anti-GnRH antibody titres. The present sequence is a synthetic
CC   immunogen of the invention
XX
SQ   Sequence 50 AA;

  Query Match             53.7%;  Score 87;  DB 5;  Length 50;
  Best Local Similarity   100.0%;  Pred. No. 3.7e-05;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 HWSYGLRPGSSGPSL 16
              ||||||||||||||||
Db          2 HWSYGLRPGSSGPSL 16


RESULT 9
AAB20147
ID   AAB20147 standard; protein; 109 AA.
XX
AC   AAB20147;
XX
DT   30-APR-2001  (first entry)
XX
DE   Growth differentiation factor 8 AutoVac construct GDF-8 P2-3.
XX
KW   Growth differentiation factor 8; GDF-8; myostatin; tetanus toxin;
KW   T-cell epitope; down-regulation; vaccine; muscle; meat; cachexia;
KW   cardiant; human; mutant; mutein.
XX
OS   Homo sapiens.
OS   Clostridium tetani.
OS   Synthetic.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   Region          1. .82
FT                   /note= "identical to residues 267-348 of human GDF-8"
FT   Misc-difference 73
FT                   /note= "Cys-73 may be substituted by Ser to avoid
FT                   disulfide bond formation"
FT   Region          83. .97
FT                   /note= "tetanus toxoid P2 epitope"
FT   Misc-difference 90. .91
FT                   /note= "optionally replaced by Glu-Gly"
FT   Region          98. .109
FT                   /note= "identical to residues 364-375 of human GDF-8"
XX
PN   WO200105820-A2.
XX
PD   25-JAN-2001.
XX
PF   20-JUL-2000; 2000WO-DK000413.
XX
PR   20-JUL-1999;   99DK-00001014.
PR   26-JUL-1999;   99US-0145275P.
XX
PA   (MEBI-) M & E BIOTECH AS.
XX
PI   Halkier T,  Mouritsen S,  Klysner S;
XX
DR   WPI; 2001-112680/12.
XX
PT   Increasing the muscle mass of animals used in meat production by down
PT   regulating growth differentiation factor 8 (GDF-8) activity in the animal
PT   through induction of anti-GDF-8 antibody production.
XX
PS   Example 1; Page 99; 110pp; English.
XX
CC   The present sequence is that of AutoVac construct GDF-8 P2-3, comprising
CC   the 109 C-terminal amino acid residues of human growth differentiation
CC   factor 8 (GDF-8) in which residues 83-97 are replaced by the promiscuous
CC   tetanus toxin T-cell epitope P2 (see AAB20143). It is an object of the
CC   invention to produce a recombinant therapeutic vaccine that is capable of
CC   effecting down-regulation of GDF-8 in order to increase the muscle growth
CC   rate of farm animals. The vaccines (see AAB20145-53) are capable of
CC   breaking autotolerance against autologous GDF-8. They comprise the C-
CC   terminal portion of human GDF-8 in which a portion of the native sequence
CC   is replaced by a T-cell epitope such as P2, with minimal disturbance of
CC   the authentic 3-dimensional structure of the protein. Nucleic acids
CC   encoding the GDF-8 variants can be used for genetic immunisation of the
CC   animals. Down-regulation of GDF-8 activity can increase muscle mass by up
CC   to at least 45% in cattle, pigs and poultry used for meat production,
CC   reducing the need for antibiotic feed-additives. Anti-GDF8 vaccines can
CC   be used to treat human diseases such as cancer cachexia where muscle
CC   atrophy is pronounced and for patients suffering from acute and chronic
CC   heart failure
XX
SQ   Sequence 109 AA;

  Query Match             52.5%;  Score 85;  DB 4;  Length 109;
  Best Local Similarity   63.3%;  Pred. No. 0.00018;
  Matches   19;  Conservative    2;  Mismatches    1;  Indels    8;  Gaps    1;

Qy         10 GSSGP--------SLQYIKANSKFIGITEL 31
              ||:||         :|||||||||||||||
Db         68 GSAGPCCTPTKMSPIQYIKANSKFIGITEL 97


RESULT 10
ABP51515
ID   ABP51515 standard; peptide; 24 AA.
XX
AC   ABP51515;
XX
DT   11-SEP-2002  (first entry)
XX
DE   HBV antigen associated peptide #8.
XX
KW   Hepatitis B virus; HBV; antigen; major histocompatibility complex; MHC;
KW   cytotoxic T cell; helper T cell; virucide; hepatotropic; immunogenic;
KW   cytotoxic T lymphocyte; CTL; HLA-restricted response.
XX
OS   Synthetic.
XX
PN   US6322789-B1.
XX
PD   27-NOV-2001.
XX
PF   05-JUN-1995;   95US-00464496.
XX
PR   26-AUG-1991;   91US-00749568.
PR   29-JAN-1992;   92US-00827682.
PR   27-APR-1992;   92US-00874491.
PR   26-AUG-1992;   92US-00935811.
XX
PA   (EPIM-) EPIMMUNE INC.
XX
PI   Vitiello MA,  Chesnut RW;
XX
DR   WPI; 2002-497942/53.
XX
PT   Immunogenic compositions for protecting against hepatitis B virus
PT   infection.
XX
PS   Claim 22; Col 39-40; 49pp; English.
XX
CC   The invention relates to a novel immunogenic composition comprising a
CC   peptide that binds to an Major Histocompatibility Complex (MHC) class I
CC   molecule to form a complex recognised by a cytotoxic T cell, and a second
CC   peptide that binds to an MHC class II molecule to form a complex
CC   recognised by a helper T cell (a group in the first peptide comprises a
CC   hepatitis B virus group). The composition of the invention has virucide
CC   and hepatotropic activity. The cytotoxic T lymphocyte (CTL)-stimulating
CC   peptides induce HLA-restricted responses to hepatitis B virus (HBV)
CC   antigens. The peptides, derived from CTL group regions of both HBV
```

```
CC   surface and nucleocapsid antigens, are particularly useful in the
CC   treatment and prevention of HBV infection, including the treatment of
CC   chronically infected HBV carriers. The peptides are also useful in
CC   diagnostic methods, such as predicting which HBV-infected individuals are
CC   prone to developing chronic infection. The sequences shown in ABP51485-
CC   ABP5559 are peptides used for the production of the immunogenic
CC   composition of the invention
XX
SQ   Sequence 24 AA;

  Query Match             51.9%;  Score 84;  DB 5;  Length 24;
  Best Local Similarity   77.3%;  Pred. No. 4.4e-05;
  Matches   17;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          9 PGSSGPSLQYIKANSKFIGITE 30
              |    ||:|||||||||||||||
Db          3 PSDFFPSVQYIKANSKFIGITE 24


RESULT 11
AAY92665
ID   AAY92665 standard; peptide; 216 AA.
XX
AC   AAY92665;
XX
DT   10-AUG-2000  (first entry)
XX
DE   MUC-1 analogue containing foreign epitopes.
XX
KW   Mucin repeat; MUC-1 analogue; vaccination; self-protein; cancer;
KW   cytotoxic T-lymphocyte immunity; breast cancer; prostate cancer;
KW   cell-associated peptide antigen; foreign epitope.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         61. .75
FT                   /label= P2
FT   Peptide         136. .156
FT                   /label= P30
FT                   /note= "q"
XX
PN   WO200020027-A2.
XX
PD   13-APR-2000.
XX
PF   05-OCT-1999;   99WO-DK000525.
XX
PR   05-OCT-1998;   98DK-00001261.
PR   20-OCT-1998;   98US-0105011P.
XX
PA   (MEBI-) M & E BIOTECH AS.
XX
PI   Steinaa L,  Mouritsen S,  Nielsen KG,  Haaning J,  Leach D,  Dalum I;
PI   Gautam A,  Birk P,  Karlsson G;
XX
DR   WPI; 2000-349917/30.
XX
PT   Inducing immune responses to weakly immunogenic, tumor associated peptide
PT   antigens for the treatment of breast and prostate cancer.
XX
PS   Example 4; Page; 220pp; English.
XX
CC   This is an immunogenized MUC-1 analogue containing foreign epitopes P2
CC   and P30. Immunogenic analogues of MUC-1 and, e.g. human prostate specific
CC   membrane antigen (hPSM) can be used in the claimed method as an
CC   autovaccine to induce a CTL response. Subdominant CTL epitopes, antibody
CC   binding regions and cysteine residues involved in disulfide bonds are
CC   preserved in the immunogenized forms (see features table). 10 regions
CC   suitable for the insertion of foreign T helper epitopes were identified.
CC   The method is used for inducing immune responses against weakly
CC   immunogenic cell-associated peptide antigens (PA) such as those
CC   associated with cancers (self-proteins), e.g. hPSM, heregulin 2 (Her2)
CC   and/or fibroblast growth factor 8b (FGF8b). The method comprises
CC   effecting simultaneous presentation by antigen producing cells (APCs) of
CC   the animals immune system of: (1) at least 1 CTL (cytotoxic T-lymphocyte)
CC   group derived from the PA and/or at least 1 B-cell group derived from the
CC   cell-associated PA; and (2) at least 1 first T helper cell group which is
CC   foreign to the animal. Analogues of human PSM, human Her2 and
CC   human/murine FGF8b comprising a substantial part of all known and
CC   predicted CTL and B-cell epitopes of the respective PA and including at
CC   least one foreign T helper epitope are also claimed. The method is used
CC   to treat prostate, prostate/breast or breast cancer when the PA is human
CC   PSM, FGF8b and Her2, respectively. Note: This sequence does not appear in
CC   the specification. It was made using the mucin repeat sequence
CC   (AAY92664), P2 and P30 (AAY92625-26), which appear on pages 220, 213 and
CC   214 respectively, of the specification
XX
SQ   Sequence 216 AA;

  Query Match             51.2%;  Score 83;  DB 3;  Length 216;
  Best Local Similarity   57.6%;  Pred. No. 0.00078;
  Matches   19;  Conservative    2;  Mismatches    2;  Indels   10;  Gaps    1;

Qy          9 PGSSGP----------SLQYIKANSKFIGITEL 31
              |||: |          : |||||||||||||||
Db         43 PGSTAPPAHGVTSAPDTRQYIKANSKFIGITEL 75


RESULT 12
AAW81331
ID   AAW81331 standard; protein; 158 AA.
XX
AC   AAW81331;
XX
DT   21-APR-1999  (first entry)
XX
DE   TNF2-7, a TNF-alpha analogue.
XX
KW   Human tumour necrosis factor-alpha; TNF-alpha; TNF-alpha analogue;
KW   vaccine; rheumatoid arthritis; Crohn's disease; ulcerative colitis;
KW   cancer; disseminated sclerosis; diabetes; psoriasis; osteoporosis;
KW   asthma.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9846642-A1.
XX
PD   22-OCT-1998.
XX
PF   15-APR-1998;   98WO-DK000157.
XX
PR   15-APR-1997;   97DK-00000418.
PR   24-APR-1997;   97US-0044187P.
XX
PA   (FERR ) FARM LAB FERRING AS.
XX
PI   Jensen MR,  Mouritsen S,  Elsner H,  Dalum I;
XX
DR   WPI; 1998-594561/50.
DR   N-PSDB; AAV68420.
XX
PT   Modified human tumour necrosis factor-alpha - comprises immunodominant T
PT   cell epitope, useful in vaccines to treat or prevent TNF-associated
PT   diseases, e.g. cancer.
XX
PS   Claim 13; Page 73; 134pp; English.
XX
CC   The present sequence represents a modified human tumour necrosis factor-
CC   alpha (TNF-alpha) analogue. The analogues have no residual TNF activity
CC   and are immunogenic in a large proportion of the human population (by
CC   using promiscuous epitopes). The TNF-alpha analogue is able to generate,
CC   in humans, neutralizing antibodies to wild-type human TNF alpha, has at
CC   least one fragment of TNF substituted by a peptide containing an
CC   immunodominant T-cell epitope, and at least one TNF-alpha B-cell epitope.
```

CC The substitution causes a significant change in the amino acid sequence
CC of any one of the strands in the front beta-sheet, any of the connecting
CC loops or any of the B', I or D strands in the back beta-sheet. The TNF-
CC alpha analogues are used as vaccines for treatment or prevention of
CC diseases associated with excessive release or activity of TNF-alpha, e.g.
CC rheumatoid arthritis, Crohn's disease, ulcerative colitis, cancer of any
CC sort, disseminated sclerosis, diabetes, psoriasis, osteoporosis and
CC asthma
XX
SQ Sequence 158 AA;

Query Match 50.9%; Score 82.5; DB 2; Length 158;
Best Local Similarity 65.6%; Pred. No. 0.00064;
Matches 21; Conservative 1; Mismatches 7; Indels 3; Gaps 1;

```
Qy          3 WSYGLRPGSSGPS---LQYIKANSKFIGITEL 31
              :|  |   |  ||   |||||||||||||||
Db         60 YSQVLFKGQGCPSTHVLQYIKANSKFIGITEL 91
```

RESULT 13
ABB07277
ID ABB07277 standard; protein; 158 AA.
XX
AC ABB07277;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human TNF-alpha analogue TNF2-7.
XX
KW TNF-alpha; pharmaceutical; vaccine; self-protein; tumour necrosis factor;
KW cetylpyridinium chloride; immunisation; antiinflammatory; antirheumatic;
KW antiarthritic; antiulcer; cytostatic; antidiabetic; antipsoriatic;
KW antiasthmatic; immunomodulator; neuroprotective; osteopathic; human;
KW TNF2-7.
XX
OS Homo sapiens.
XX
PN WO200197837-A1.
XX
PD 27-DEC-2001.
XX
PF 20-JUN-2001; 2001WO-DK000431.
XX
PR 21-JUN-2000; 2000DK-00000966.
XX
PA (FERR) FERRING BV.
XX
PI Olesen OF, Balchen T, Bouman MHEM;
XX
DR WPI; 2002-114542/15.
DR N-PSDB; ABA94387.
XX
PT Novel vaccine composition for prevention/treatment of self-protein-
PT mediated pathology such as cancer, diabetes and asthma, comprises
PT modified immunogenic self-protein and surfactant capable of acting as
PT solubilizer.
XX
PS Claim 21; Page 39; 55pp; English.
XX
CC The invention provides a pharmaceutical vaccine composition (I) for the
CC prevention or treatment of a self-protein-mediated pathology. The
CC composition comprises at least one modified immunogenic self-protein
CC (selected from modified TNF-alpha proteins) and a surfactant capable of
CC acting as a solubilizer. (I) is useful for preventing or treating a self
CC -protein-mediated pathology such as an inflammatory disease, rheumatoid
CC arthritis, an inflammatory bowel disease (ulcerative colitis or Crohn's
CC disease), cancer, cachexia, multiple sclerosis, diabetes, psoriasis,
CC osteoporosis or asthma. (I) is useful for inducing autoantibodies to a
CC self-protein such as TNF (tumour necrosis factor)-alpha in a human
CC subject. (I) comprising cetylpyridinium chloride as a component is useful
CC for immunisation of a human subject or for treatment of a human
CC inflammatory disease. The present sequence represents a human TNF-alpha
CC analogue TNF2-7
XX
SQ Sequence 158 AA;

Query Match 50.9%; Score 82.5; DB 5; Length 158;
Best Local Similarity 65.6%; Pred. No. 0.00064;
Matches 21; Conservative 1; Mismatches 7; Indels 3; Gaps 1;

```
Qy          3 WSYGLRPGSSGPS---LQYIKANSKFIGITEL 31
              :|  |   |  ||   |||||||||||||||
Db         60 YSQVLFKGQGCPSTHVLQYIKANSKFIGITEL 91
```

RESULT 14
AAB45502
ID AAB45502 standard; protein; 116 AA.
XX
AC AAB45502;
XX
DT 26-FEB-2001 (first entry)
XX
DE Modified murine interleukin-5 SEQ ID NO: 14.
XX
KW Asthma; IL-5; interleukin-5; allergy; cytokine; helminthic infection;
KW cancer; eosinophilia; vaccine; allergic rhinitis.
XX
OS Mus musculus.
OS Clostridium tetani.
XX
PN WO200065058-A1.
XX
PD 02-NOV-2000.
XX
PF 19-APR-2000; 2000WO-DK000205.
XX
PR 23-APR-1999; 99DK-00000552.
PR 06-MAY-1999; 99US-0132811P.
XX
PA (MEBI-) M & E BIOTECH AS.
XX
PI Klysner S;
XX
DR WPI; 2000-672791/65.
XX
PT Down-regulating interleukin 5 (IL-5) activity in humans by administering
PT IL-5 and/or an IL-5 analogue, useful in the treatment, prophylaxis or
PT amelioration of asthma or other chronic allergic conditions.
XX
PS Example 2; Page 129-130; 172pp; English.
XX
CC The present invention is concerned with methods of treating asthma,
CC eosinophilia, allergic rhinitis and other allergic diseases. These
CC involve the use of interleukin-5 (IL-5) analogues and modified IL-5
CC proteins and their coding sequences to down-regulate IL-5 activity and
CC thus reduce eosinophil numbers. The allergic diseases may be treated
CC using autovaccines, nucleic acid vaccines or live vaccines. In addition,
CC it is possible that they may be used in the treatment of cancer and
CC helminthic infections
XX
SQ Sequence 116 AA;

Query Match 50.0%; Score 81; DB 3; Length 116;
Best Local Similarity 76.2%; Pred. No. 0.00074;
Matches 16; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

```
Qy         11 SSGPSLQYIKANSKFIGITEL 31
              :|  ::||||||||||||||||
Db         24 TSNETMQYIKANSKFIGITEL 44
```

RESULT 15
AAB45526
ID AAB45526 standard; protein; 116 AA.

XX
AC AAB45526;
XX
DT 26-FEB-2001 (first entry)
XX
DE Modified murine interleukin-5 SEQ ID NO: 52.
XX
KW Asthma; IL-5; interleukin-5; allergy; cytokine; helminthic infection;
KW cancer; eosinophilia; vaccine; allergic rhinitis.
XX
OS Mus musculus.
OS Clostridium tetani.
XX
PN WO200065058-A1.
XX
PD 02-NOV-2000.
XX
PF 19-APR-2000; 2000WO-DK000205.
XX
PR 23-APR-1999; 99DK-00000552.
PR 06-MAY-1999; 99US-0132811P.
XX
PA (MEBI-) M & E BIOTECH AS.
XX
PI Klysner S;
XX
DR WPI; 2000-672791/65.
DR N-PSDB; AAC68879.
XX
PT Down-regulating interleukin 5 (IL-5) activity in humans by administering
PT IL-5 and/or an IL-5 analogue, useful in the treatment, prophylaxis or
PT amelioration of asthma or other chronic allergic conditions.
XX
PS Disclosure; Page 159-160; 172pp; English.
XX
CC The present invention is concerned with methods of treating asthma,
CC eosinophilia, allergic rhinitis and other allergic diseases. These
CC involve the use of interleukin-5 (IL-5) analogues and modified IL-5
CC proteins and their coding sequences to down-regulate IL-5 activity and
CC thus reduce eosinophil numbers. The allergic diseases may be treated
CC using autovaccines, nucleic acid vaccines or live vaccines. In addition,
CC it is possible that they may be used in the treatment of cancer and
CC helminthic infections
XX
SQ Sequence 116 AA;

Query Match 50.0%; Score 81; DB 3; Length 116;
Best Local Similarity 76.2%; Pred. No. 0.00074;
Matches 16; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

```
Qy         11 SSGPSLQYIKANSKFIGITEL 31
              :|  ::|||||||||||||||
Db         24 TSNETMQYIKANSKFIGITEL 44
```

Search completed: March 10, 2004, 09:12:12
Job time : 47.6809 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 10, 2004, 09:16:59 ; Search time 24.3658 Seconds
(without alignments)
268.645 Million cell updates/sec

Title: US-09-848-834A-15
Perfect score: 162
Sequence: 1 XHWSYGLRPGSSGPSLQYIKANSKFIGITEL 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	161	99.4	31	9	US-09-848-834A-15	Sequence 15, Appl
2	161	99.4	46	9	US-09-848-834A-19	Sequence 19, Appl
3	89	54.9	34	9	US-09-848-834A-13	Sequence 13, Appl
4	89	54.9	36	9	US-09-848-834A-16	Sequence 16, Appl
5	89	54.9	47	9	US-09-848-834A-17	Sequence 17, Appl
6	89	54.9	51	9	US-09-848-834A-20	Sequence 20, Appl
7	87	53.7	37	9	US-09-848-834A-14	Sequence 14, Appl
8	87	53.7	50	9	US-09-848-834A-18	Sequence 18, Appl
9	82.5	50.9	158	14	US-10-297-942-8	Sequence 8, Appli
10	80.5	49.7	158	14	US-10-297-942-14	Sequence 14, Appl
11	79	48.8	287	14	US-10-295-074-15	Sequence 15, Appl
12	77	47.5	194	14	US-10-295-074-46	Sequence 46, Appl
13	76	46.9	16	9	US-09-848-834A-2	Sequence 2, Appli
14	76	46.9	17	15	US-10-346-563-23	Sequence 23, Appl
15	76	46.9	17	15	US-10-321-717-23	Sequence 23, Appl
16	76	46.9	194	14	US-10-295-074-47	Sequence 47, Appl
17	76	46.9	285	14	US-10-295-074-9	Sequence 9, Appli
18	76	46.9	573	15	US-10-452-024-177	Sequence 177, App
19	76	46.9	872	14	US-10-241-596-145	Sequence 145, App
20	76	46.9	879	14	US-10-241-596-143	Sequence 143, App
21	76	46.9	887	14	US-10-241-596-147	Sequence 147, App
22	76	46.9	1310	15	US-10-452-024-149	Sequence 149, App
23	76	46.9	1315	14	US-10-241-596-141	Sequence 141, App
24	76	46.9	1315	15	US-10-452-024-145	Sequence 145, App
25	74	45.7	15	9	US-09-862-849-2	Sequence 2, Appli
26	74	45.7	15	9	US-09-785-215-4	Sequence 4, Appli
27	74	45.7	15	10	US-09-405-986-1	Sequence 1, Appli
28	74	45.7	15	14	US-10-204-362-4	Sequence 4, Appli
29	74	45.7	15	14	US-10-223-711-7	Sequence 7, Appli
30	74	45.7	15	14	US-10-237-656-13	Sequence 13, Appl
31	74	45.7	15	14	US-10-223-809A-4	Sequence 4, Appli
32	74	45.7	15	14	US-10-261-446-19	Sequence 19, Appl
33	74	45.7	15	14	US-10-239-313A-618	Sequence 618, App
34	74	45.7	15	14	US-10-295-074-3	Sequence 3, Appli
35	74	45.7	15	15	US-10-372-111-7	Sequence 7, Appli
36	74	45.7	16	10	US-09-930-915A-64	Sequence 64, Appl
37	74	45.7	16	14	US-10-082-014-145	Sequence 145, App
38	74	45.7	16	14	US-10-372-076-175	Sequence 175, App
39	74	45.7	17	10	US-09-865-294-3	Sequence 3, Appli
40	74	45.7	17	14	US-10-239-313A-619	Sequence 619, App
41	74	45.7	17	15	US-10-411-544-5	Sequence 5, Appli
42	74	45.7	19	14	US-10-239-313A-620	Sequence 620, App
43	74	45.7	27	14	US-10-076-674-7	Sequence 7, Appli
44	74	45.7	27	15	US-10-355-161A-7	Sequence 7, Appli
45	74	45.7	28	9	US-09-848-834A-11	Sequence 11, Appl

ALIGNMENTS

RESULT 1
US-09-848-834A-15
; Sequence 15, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Aphton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the
; OTHER INFORMATION: RH hormone linked by a spacer to amino sequence 830-844 of the
; OTHER INFORMATION: tanus toxoid precursor (Tentoxylysin)
; NAME/KEY: MOD_RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(10)
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
; NAME/KEY: PEPTIDE
; LOCATION: (11)..(16)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (17)..(31)
; OTHER INFORMATION: Amino acid sequence 830-844 of the Tetanus toxoid precursor
; OTHER INFORMATION: (Tentoxylysin)
US-09-848-834A-15

Query Match 99.4%; Score 161; DB 9; Length 31;

```
  Best Local Similarity   100.0%;  Pred. No. 5.3e-16;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 HWSYGLRPGSSGPSLQYIKANSKFIGITEL 31
              |||||||||||||||||||||||||||||||
Db          2 HWSYGLRPGSSGPSLQYIKANSKFIGITEL 31


RESULT 2
US-09-848-834A-19
; Sequence 19, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
;  APPLICANT: Aphton Corporation
;  TITLE OF INVENTION: Chimeric Peptide Immunogens
;  FILE REFERENCE: 1102865-0047
;  CURRENT APPLICATION NUMBER: US/09/848,834A
;  CURRENT FILING DATE:  2001-05-04
;  PRIOR APPLICATION NUMBER: 60/202,328
;  PRIOR FILING DATE: 2000-05-05
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 19
;   LENGTH: 46
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of human
;   OTHER INFORMATION: GnRH linked by a spacer to amino acid sequence 830-844 of Tetanus
;   OTHER INFORMATION: toxoid precursor (Tentoxylysin) linked by a spacer to amino acid
;   OTHER INFORMATION: sequence 1-10 of GnRH
;   NAME/KEY: MOD_RES
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
;   NAME/KEY: MOD_RES
;   LOCATION: (46)..(46)
;   OTHER INFORMATION: Amidated glycine or glycinamide
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)..(10)
;   OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
;   NAME/KEY: PEPTIDE
;   LOCATION: (11)..(16)
;   OTHER INFORMATION: Spacer peptide
;   NAME/KEY: PEPTIDE
;   LOCATION: (17)..(31)
;   OTHER INFORMATION: Amino acid sequence 830-844 of the Tetanus toxoid precursor
;   OTHER INFORMATION: (Tentoxylysin)
;   NAME/KEY: PEPTIDE
;   LOCATION: (32)..(37)
;   OTHER INFORMATION: Spacer peptide
;   NAME/KEY: PEPTIDE
;   LOCATION: (38)..(46)
;   OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone
US-09-848-834A-19

  Query Match             99.4%;  Score 161;  DB 9;  Length 46;
  Best Local Similarity   100.0%;  Pred. No. 8.2e-16;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 HWSYGLRPGSSGPSLQYIKANSKFIGITEL 31
              |||||||||||||||||||||||||||||||
Db          2 HWSYGLRPGSSGPSLQYIKANSKFIGITEL 31


RESULT 3
US-09-848-834A-13
; Sequence 13, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
;  APPLICANT: Aphton Corporation
;  TITLE OF INVENTION: Chimeric Peptide Immunogens
;  FILE REFERENCE: 1102865-0047
;  CURRENT APPLICATION NUMBER: US/09/848,834A
;  CURRENT FILING DATE:  2001-05-04
;  PRIOR APPLICATION NUMBER: 60/202,328
;  PRIOR FILING DATE: 2000-05-05
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 13
;   LENGTH: 34
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the G
;   OTHER INFORMATION: RH hormone linked by a spacer to amino acid sequence 288-302 of
;   OTHER INFORMATION: he Measles virus fusion protein,
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)..(10)
;   OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
;   NAME/KEY: PEPTIDE
;   LOCATION: (11)..(18)
;   OTHER INFORMATION: Spacer peptide
;   NAME/KEY: PEPTIDE
;   LOCATION: (19)..(34)
;   OTHER INFORMATION: Amino acid sequence 288-302 of the Measles
;   OTHER INFORMATION: virus fusion protein, F
;   NAME/KEY: MOD_RES
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
US-09-848-834A-13

  Query Match             54.9%;  Score 89;  DB 9;  Length 34;
  Best Local Similarity   83.3%;  Pred. No. 1.3e-05;
  Matches   15;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 HWSYGLRPGSSGPSLQYI 19
              |||||||||||||||: :
Db          2 HWSYGLRPGSSGPSLKLL 19


RESULT 4
US-09-848-834A-16
; Sequence 16, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
;  APPLICANT: Aphton Corporation
;  TITLE OF INVENTION: Chimeric Peptide Immunogens
;  FILE REFERENCE: 1102865-0047
;  CURRENT APPLICATION NUMBER: US/09/848,834A
;  CURRENT FILING DATE:  2001-05-04
;  PRIOR APPLICATION NUMBER: 60/202,328
;  PRIOR FILING DATE: 2000-05-05
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 16
;   LENGTH: 36
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the G
;   OTHER INFORMATION: RH hormone linked by a spacer to amino acid sequence 378-398 of
;   OTHER INFORMATION: the Plasmodium falciparum circumsporozoite (CSP) protein
;   NAME/KEY: MOD_RES
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)..(10)
;   OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
;   NAME/KEY: PEPTIDE
;   LOCATION: (11)..(16)
;   OTHER INFORMATION: Spacer peptide
;   NAME/KEY: PEPTIDE
;   LOCATION: (17)..(36)
;   OTHER INFORMATION: Amino acid sequence 378-398 of the Malaria
;   OTHER INFORMATION: (Plasmodium falciparum) circumsporozoite
;   OTHER INFORMATION: (CSP) protein
```

```
US-09-848-834A-16

  Query Match             54.9%;  Score 89;  DB 9;  Length 36;
  Best Local Similarity   63.3%;  Pred. No. 1.3e-05;
  Matches   19;  Conservative    2;  Mismatches    3;  Indels    6;  Gaps    1;

Qy          2 HWSYGLRPGSSGPSL------QYIKANSKF 25
              |||||||||||||||      :  ||:| |
Db          2 HWSYGLRPGSSGPSLDEKKIAKMEKASSVF 31


RESULT 5
US-09-848-834A-17
; Sequence 17, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
;  APPLICANT: Aphton Corporation
;  TITLE OF INVENTION: Chimeric Peptide Immunogens
;  FILE REFERENCE: 1102865-0047
;  CURRENT APPLICATION NUMBER: US/09/848,834A
;  CURRENT FILING DATE:  2001-05-04
;  PRIOR APPLICATION NUMBER: 60/202,328
;  PRIOR FILING DATE: 2000-05-05
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
;   LENGTH: 47
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the Gn
;   OTHER INFORMATION: RH hormone linked by a spacer to amino acid sequence 288-302 of
;   OTHER INFORMATION: the Measles virus protein F linked by a spacer to amino acid seq
;   OTHER INFORMATION: uence 2-10 of the GnRH hormone
;   NAME/KEY: MOD_RES
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
;   NAME/KEY: MOD_RES
;   LOCATION: (47)..(47)
;   OTHER INFORMATION: Amidated-glycine or glycinamide
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)..(10)
;   OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
;   NAME/KEY: PEPTIDE
;   LOCATION: (11)..(18)
;   OTHER INFORMATION: Spacer peptide
;   NAME/KEY: PEPTIDE
;   LOCATION: (19)..(34)
;   OTHER INFORMATION: Amino acid sequence 288-302 of the Measles virus fusion protein,
;   NAME/KEY: PEPTIDE
;   LOCATION: (35)..(38)
;   OTHER INFORMATION: Spacer peptide
;   NAME/KEY: PEPTIDE
;   LOCATION: (39)..(47)
;   OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone
US-09-848-834A-17

  Query Match             54.9%;  Score 89;  DB 9;  Length 47;
  Best Local Similarity   83.3%;  Pred. No. 1.8e-05;
  Matches   15;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 HWSYGLRPGSSGPSLQYI 19
              |||||||||||||||: :
Db          2 HWSYGLRPGSSGPSLKLL 19


RESULT 6
US-09-848-834A-20
; Sequence 20, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
;  APPLICANT: Aphton Corporation
;  TITLE OF INVENTION: Chimeric Peptide Immunogens
;  FILE REFERENCE: 1102865-0047
;  CURRENT APPLICATION NUMBER: US/09/848,834A
;  CURRENT FILING DATE:  2001-05-04
;  PRIOR APPLICATION NUMBER: 60/202,328
;  PRIOR FILING DATE: 2000-05-05
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
;   LENGTH: 51
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of huma
;   OTHER INFORMATION: GnRH linked by a spacer to amino acid sequence 378-398 of Plasm
;   OTHER INFORMATION: ium falciparum circumsporozoite (CSP) protein
;   NAME/KEY: MOD_RES
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
;   NAME/KEY: MOD_RES
;   LOCATION: (51)..(51)
;   OTHER INFORMATION: Amidated glycine or glycinamide
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)..(10)
;   OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
;   NAME/KEY: PEPTIDE
;   LOCATION: (11)..(16)
;   OTHER INFORMATION: Spacer peptide
;   NAME/KEY: PEPTIDE
;   LOCATION: (17)..(36)
;   OTHER INFORMATION: Amino acid sequence 378-398 of the Plasmodium falciparum
;   OTHER INFORMATION: circumsporozoite (CSP) protein
;   NAME/KEY: PEPTIDE
;   LOCATION: (37)..(42)
;   OTHER INFORMATION: Spacer peptide
;   NAME/KEY: PEPTIDE
;   LOCATION: (43)..(51)
;   OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone
US-09-848-834A-20

  Query Match             54.9%;  Score 89;  DB 9;  Length 51;
  Best Local Similarity   63.3%;  Pred. No. 2e-05;
  Matches   19;  Conservative    2;  Mismatches    3;  Indels    6;  Gaps    1;

Qy          2 HWSYGLRPGSSGPSL------QYIKANSKF 25
              |||||||||||||||      :  ||:| |
Db          2 HWSYGLRPGSSGPSLDEKKIAKMEKASSVF 31


RESULT 7
US-09-848-834A-14
; Sequence 14, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
;  APPLICANT: Aphton Corporation
;  TITLE OF INVENTION: Chimeric Peptide Immunogens
;  FILE REFERENCE: 1102865-0047
;  CURRENT APPLICATION NUMBER: US/09/848,834A
;  CURRENT FILING DATE:  2001-05-04
;  PRIOR APPLICATION NUMBER: 60/202,328
;  PRIOR FILING DATE: 2000-05-05
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
;   LENGTH: 37
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the
;   OTHER INFORMATION: nRH hormone linked by a spacer to amino acid sequence 947-967 c
;   OTHER INFORMATION: the Tetanus toxoid precursor (Tentoxylysin)
;   NAME/KEY: MOD_RES
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
```

```
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)..(10)
;   OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
;   NAME/KEY: PEPTIDE
;   LOCATION: (11)..(16)
;   OTHER INFORMATION: Spacer peptide
;   NAME/KEY: PEPTIDE
;   LOCATION: (17)..(37)
;   OTHER INFORMATION: Amino acid sequence 947-967 of the Tetanus toxoid precursor
;   OTHER INFORMATION: (Tentoxylysin)
US-09-848-834A-14

  Query Match             53.7%;  Score 87;  DB 9;  Length 37;
  Best Local Similarity   100.0%;  Pred. No. 2.7e-05;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 HWSYGLRPGSSGPSL 16
              ||||||||||||||||
Db          2 HWSYGLRPGSSGPSL 16


RESULT 8
US-09-848-834A-18
; Sequence 18, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
;  APPLICANT: Aphton Corporation
;  TITLE OF INVENTION: Chimeric Peptide Immunogens
;  FILE REFERENCE: 1102865-0047
;  CURRENT APPLICATION NUMBER: US/09/848,834A
;  CURRENT FILING DATE:  2001-05-04
;  PRIOR APPLICATION NUMBER: 60/202,328
;  PRIOR FILING DATE: 2000-05-05
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18
;   LENGTH: 50
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of human
;   OTHER INFORMATION: GnRH linked by a spacer to amino acid sequence 947-967 of the Tet
;   OTHER INFORMATION: anus toxoid precursor (Tentoxylysin) protein linked by a spacer t
;   OTHER INFORMATION: o amino acid sequence 2-10 of human GnRH
;   NAME/KEY: MOD_RES
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
;   NAME/KEY: MOD_RES
;   LOCATION: (50)..(50)
;   OTHER INFORMATION: Amidated glycine or glycinamide
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)..(10)
;   OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
;   NAME/KEY: PEPTIDE
;   LOCATION: (11)..(16)
;   OTHER INFORMATION: Spacer peptide
;   NAME/KEY: PEPTIDE
;   LOCATION: (17)..(37)
;   OTHER INFORMATION: Amino acid sequence 947-967 of the Tetanus toxoid precursor (Tent
;   OTHER INFORMATION: oxylysin
;   NAME/KEY: PEPTIDE
;   LOCATION: (38)..(41)
;   OTHER INFORMATION: Spacer peptide
;   NAME/KEY: PEPTIDE
;   LOCATION: (42)..(50)
;   OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone
US-09-848-834A-18

  Query Match             53.7%;  Score 87;  DB 9;  Length 50;
  Best Local Similarity   100.0%;  Pred. No. 3.7e-05;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 HWSYGLRPGSSGPSL 16
              ||||||||||||||||
Db          2 HWSYGLRPGSSGPSL 16


RESULT 9
US-10-297-942-8
; Sequence 8, Application US/10297942
; Publication No. US20030185816A1
; GENERAL INFORMATION:
;  APPLICANT: Ferring BV
;  TITLE OF INVENTION: Solubilised Protein Vaccines
;  FILE REFERENCE: P68445US0
;  CURRENT APPLICATION NUMBER: US/10/297,942
;  CURRENT FILING DATE:  2003-04-21
;  PRIOR APPLICATION NUMBER: PCT/DK01/00431
;  PRIOR FILING DATE: 2001-10-16
;  PRIOR APPLICATION NUMBER: DK PA 2000 00966
;  PRIOR FILING DATE: 2000-06-21
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
;   LENGTH: 158
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-297-942-8

  Query Match             50.9%;  Score 82.5;  DB 14;  Length 158;
  Best Local Similarity   65.6%;  Pred. No. 0.00058;
  Matches   21;  Conservative    1;  Mismatches    7;  Indels    3;  Gaps    1;

Qy          3 WSYGLRPGSSGPS---LQYIKANSKFIGITEL 31
              :|  |  |   ||   ||||||||||||||||
Db         60 YSQVLFKGQGCPSTHVLQYIKANSKFIGITEL 91


RESULT 10
US-10-297-942-14
; Sequence 14, Application US/10297942
; Publication No. US20030185816A1
; GENERAL INFORMATION:
;  APPLICANT: Ferring BV
;  TITLE OF INVENTION: Solubilised Protein Vaccines
;  FILE REFERENCE: P68445US0
;  CURRENT APPLICATION NUMBER: US/10/297,942
;  CURRENT FILING DATE:  2003-04-21
;  PRIOR APPLICATION NUMBER: PCT/DK01/00431
;  PRIOR FILING DATE: 2001-10-16
;  PRIOR APPLICATION NUMBER: DK PA 2000 00966
;  PRIOR FILING DATE: 2000-06-21
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
;   LENGTH: 158
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-297-942-14

  Query Match             49.7%;  Score 80.5;  DB 14;  Length 158;
  Best Local Similarity   76.0%;  Pred. No. 0.0011;
  Matches   19;  Conservative    1;  Mismatches    4;  Indels    1;  Gaps    1;

Qy          7 LRPGSSGPSLQYIKANSKFIGITEL 31
              :|  |  || ||||||||||||||||
Db          2 VRSSSRTPS-QYIKANSKFIGITEL 25


RESULT 11
US-10-295-074-15
; Sequence 15, Application US/10295074
; Publication No. US20030185845A1
; GENERAL INFORMATION:
;  APPLICANT: Pharmexa A/S
;  TITLE OF INVENTION: NOVEL IMMUNOGENIC MIMETICS OF MULTIMER PROTEINS
```

```
;  FILE REFERENCE: P1013DK00
;  CURRENT APPLICATION NUMBER: US/10/295,074
;  CURRENT FILING DATE:  2002-11-15
;  NUMBER OF SEQ ID NOS: 60
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
;   LENGTH: 287
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: Two human IL5 monomers joined by a di-glycine linker and includin
;   OTHER INFORMATION: g terminally positioned tetanus toxoid P2 and P30 epitopes
US-10-295-074-15

  Query Match             48.8%;  Score 79;  DB 14;  Length 287;
  Best Local Similarity   88.9%;  Pred. No. 0.0035;
  Matches   16;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         14 PSLQYIKANSKFIGITEL 31
              |: ||||||||||||||||
Db         21 PTEQYIKANSKFIGITEL 38


RESULT 12
US-10-295-074-46
; Sequence 46, Application US/10295074
; Publication No. US20030185845A1
; GENERAL INFORMATION:
;  APPLICANT: Pharmexa A/S
;  TITLE OF INVENTION: NOVEL IMMUNOGENIC MIMETICS OF MULTIMER PROTEINS
;  FILE REFERENCE: P1013DK00
;  CURRENT APPLICATION NUMBER: US/10/295,074
;  CURRENT FILING DATE:  2002-11-15
;  NUMBER OF SEQ ID NOS: 60
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 46
;   LENGTH: 194
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: hTNF with inserted tetanus toxoid P2 and P30 epitopes
;   FEATURE:
;   NAME/KEY: MUTAGEN
;   LOCATION: (110)..(124)
;   OTHER INFORMATION: Tetanus toxoid P2 epitope (SEQ ID NO: 2)
;   FEATURE:
;   NAME/KEY: MUTAGEN
;   LOCATION: (125)..(145)
;   OTHER INFORMATION: Tetanus toxoid P30 epitope (SEQ ID NO: 3)
;   FEATURE:
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (2)..(109)
;   OTHER INFORMATION: hTNF amino acids 1-108
;   FEATURE:
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (146)..(194)
;   OTHER INFORMATION: hTNF amino acids 109-157
US-10-295-074-46

  Query Match             47.5%;  Score 77;  DB 14;  Length 194;
  Best Local Similarity   88.9%;  Pred. No. 0.0045;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         14 PSLQYIKANSKFIGITEL 31
              |  ||||||||||||||||
Db        107 PEGQYIKANSKFIGITEL 124


RESULT 13
US-09-848-834A-2
; Sequence 2, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
;  APPLICANT: Aphton Corporation
;  TITLE OF INVENTION: Chimeric Peptide Immunogens
;  FILE REFERENCE: 1102865-0047
;  CURRENT APPLICATION NUMBER: US/09/848,834A
;  CURRENT FILING DATE:  2001-05-04
;  PRIOR APPLICATION NUMBER: 60/202,328
;  PRIOR FILING DATE: 2000-05-05
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
;   LENGTH: 16
;   TYPE: PRT
;   ORGANISM: Tetanus bacillus
;   FEATURE:
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)..(16)
;   OTHER INFORMATION: Amino acid sequence 829-844 of the Tetanus
;   OTHER INFORMATION: Toxoid Precursor (Tentoxylysin)
US-09-848-834A-2

  Query Match             46.9%;  Score 76;  DB 9;  Length 16;
  Best Local Similarity   93.8%;  Pred. No. 0.0004;
  Matches   15;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         16 LQYIKANSKFIGITEL 31
              :|||||||||||||||
Db          1 MQYIKANSKFIGITEL 16


RESULT 14
US-10-346-563-23
; Sequence 23, Application US/10346563
; Publication No. US20030220229A1
;   GENERAL INFORMATION:
;        APPLICANT: Hickey, William F.
;                   Griffin, Ann C.
;        TITLE OF INVENTION: Proinsulin Peptide Compounds for Detecting and
;                            Treating Type I Diabetes
;        NUMBER OF SEQUENCES: 23
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: LAHIVE & COCKFIELD
;             STREET: 60 State Street, suite 510
;             CITY: Boston
;             STATE: Massachusetts
;             COUNTRY: USA
;             ZIP: 02109-1875
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: ASCII Text
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/346,563
;             FILING DATE: 16-Jan-2003
;             CLASSIFICATION: 424
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/08/472,704
;             FILING DATE: 06-Jun-1995
;             APPLICATION NUMBER: US 08/272,220
;             FILING DATE: 08-JULY-1994
;        ATTORNEY/AGENT INFORMATION:
;             NAME: DeConti, Giulio A., Jr.
;             REGISTRATION NUMBER: 31,503
;             REFERENCE/DOCKET NUMBER: DCI-092
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (617)227-7400
;             TELEFAX: (617)227-5941
;   INFORMATION FOR SEQ ID NO: 23:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 17 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
```

```
;       FRAGMENT TYPE: internal
;       SEQUENCE DESCRIPTION: SEQ ID NO: 23:
US-10-346-563-23

  Query Match             46.9%;  Score 76;  DB 15;  Length 17;
  Best Local Similarity   93.8%;  Pred. No. 0.00043;
  Matches   15;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         16 LQYIKANSKFIGITEL 31
              :||||||||||||||||
Db          2 MQYIKANSKFIGITEL 17


RESULT 15
US-10-321-717-23
; Sequence 23, Application US/10321717
; Publication No. US20040002113A1
;   GENERAL INFORMATION:
;       APPLICANT: Griffin, Ann C.
;                  Hickey, William F.
;       TITLE OF INVENTION: Detection and Treatment Methods for
;                           Type I Diabetes
;       NUMBER OF SEQUENCES: 23
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: LAHIVE & COCKFIELD
;            STREET: 60 State Street, suite 510
;            CITY: Boston
;            STATE: Massachusetts
;            COUNTRY: USA
;            ZIP: 02109-1875
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: ASCII Text
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/321,717
;            FILING DATE: 17-Dec-2002
;            CLASSIFICATION: 435
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/08/472,701
;            FILING DATE: <Unknown>
;            APPLICATION NUMBER: US 08/272,220
;            FILING DATE: 08-JULY-1994
;       ATTORNEY/AGENT INFORMATION:
;            NAME: DeConti, Giulio A., Jr.
;            REGISTRATION NUMBER: 31,503
;            REFERENCE/DOCKET NUMBER: DCI-092DV
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (617)227-7400
;            TELEFAX: (617)227-5941
;   INFORMATION FOR SEQ ID NO: 23:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 17 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: peptide
;       FRAGMENT TYPE: internal
;       SEQUENCE DESCRIPTION: SEQ ID NO: 23:
US-10-321-717-23

  Query Match             46.9%;  Score 76;  DB 15;  Length 17;
  Best Local Similarity   93.8%;  Pred. No. 0.00043;
  Matches   15;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         16 LQYIKANSKFIGITEL 31
              :||||||||||||||||
Db          2 MQYIKANSKFIGITEL 17


Search completed: March 10, 2004, 10:25:49
Job time : 24.3658 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:54 ; Search time 9.64981 Seconds
(without alignments)
309.015 Million cell updates/sec

Title: US-09-848-834A-15
Perfect score: 162
Sequence: 1 XHWSYGLRPGSSGPSLQYIKANSKFIGITEL 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	76	46.9	1315	1	BTCLTN	tentoxilysin (EC 3
2	58	35.8	10	1	RHPGG	gonadoliberin - pi
3	58	35.8	10	1	RHSHG	gonadoliberin - sh
4	58	35.8	67	2	I78541	gonadoliberin prec
5	58	35.8	89	2	I51423	gonadoliberin prec
6	58	35.8	90	1	RHMSG	gonadoliberin prec
7	58	35.8	92	1	RHHUG	gonadoliberin prec
8	58	35.8	92	1	RHRTG	gonadoliberin prec
9	56	34.6	719	2	T52510	hypothetical prote
10	55.5	34.3	66	2	S31029	gene 84 protein -
11	55	34.0	123	2	G48677	Ig heavy chain V-D
12	54	33.3	10	1	RHAQ1	gonadoliberin I -
13	54	33.3	92	2	I50644	gonadoliberin I pr
14	53.5	33.0	374	2	E95361	probable muconate
15	52	32.1	98	2	I50739	gonadotropin-relea
16	52	32.1	102	2	PH1491	Ig heavy chain V r
17	52	32.1	119	2	PH1518	Ig heavy chain V r
18	52	32.1	119	2	PH1519	Ig heavy chain V r
19	52	32.1	135	2	PH1494	Ig heavy chain V r
20	52	32.1	575	2	T06353	isocitrate lyase (
21	51	31.5	112	2	C27887	Ig kappa chain V r
22	51	31.5	115	2	S38715	Ig kappa chain V r
23	51	31.5	123	2	F48677	Ig heavy chain V-D
24	51	31.5	208	2	AG2249	hypothetical prote
25	50.5	31.2	1494	2	T14355	protein-tyrosine-p
26	50	30.9	80	1	RHID1S	gonadoliberin I pr
27	50	30.9	120	2	A49043	Ig kappa chain V r
28	50	30.9	224	2	A53143	testis-determining
29	50	30.9	249	2	A41497	36K antigen pra -
30	49.5	30.6	256	2	S74928	hypothetical prote
31	49	30.2	75	2	AI3191	hypothetical prote
32	49	30.2	108	2	E32530	Ig kappa chain V r
33	49	30.2	112	2	D27887	Ig kappa chain V r
34	49	30.2	119	2	PH1516	Ig heavy chain V r
35	49	30.2	140	2	PH1488	Ig heavy chain V r
36	49	30.2	271	2	D71103	probable homoserin
37	49	30.2	294	2	H75080	probable homoserin
38	49	30.2	315	2	A88043	protein C13A10.3 [
39	49	30.2	444	2	C65102	probable galactara
40	49	30.2	444	2	E85974	probable galactara
41	49	30.2	444	2	E91129	probable galactara
42	49	30.2	485	1	RRYC62	RNA-directed DNA p
43	49	30.2	1047	2	T49425	hypothetical prote
44	48.5	29.9	884	2	AI0424	translation initia
45	48	29.6	91	2	JC7393	medaka-type gonado

ALIGNMENTS

RESULT 1
BTCLTN
tentoxilysin (EC 3.4.24.68) precursor - Clostridium tetani
N;Alternate names: tetanus neurotoxin
C;Species: Clostridium tetani
C;Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 03-Jun-2002
C;Accession: A25689; A25757; A25194; B25194; A60759; S69348; S09364
R;Eisel, U.; Jarausch, W.; Goretzki, K.; Henschen, A.; Engels, J.; Weller, U.; Hudel,
EMBO J. 5, 2495-2502, 1986
A;Title: Tetanus toxin: primary structure, expression in E. coli, and homology with bo
A;Reference number: A25689; MUID:87053814; PMID:3536478
A;Accession: A25689
A;Molecule type: DNA
A;Residues: 1-1315 <EIS>
A;Cross-references: GB:X04436; NID:g40769; PIDN:CAA28033.1; PID:g40770
R;Fairweather, N.F.; Lyness, V.A.
Nucleic Acids Res. 14, 7809-7812, 1986
A;Title: The complete nucleotide sequence of tetanus toxin.
A;Reference number: A25757; MUID:87040747; PMID:3774547
A;Accession: A25757
A;Molecule type: DNA
A;Residues: 1-1315 <FAI>
A;Cross-references: GB:X06214; NID:g40773; PIDN:CAA29564.1; PID:g40774
A;Experimental source: strain CN3911
R;Fairweather, N.F.; Lyness, V.A.; Pickard, D.J.; Allen, G.; Thomson, R.O.
J. Bacteriol. 165, 21-27, 1986
A;Title: Cloning, nucleotide sequencing, and expression of tetanus toxin fragment C in
A;Reference number: A25194; MUID:86085672; PMID:3510187
A;Accession: A25194
A;Molecule type: DNA
A;Residues: 743-1315 <FA2>
A;Cross-references: GB:M12739; NID:g144920; PIDN:AAA23282.1; PID:g144921
A;Accession: B25194
A;Molecule type: protein
A;Residues: 865-894 <FA3>
R;Matsuda, M.; Lei, D.L.; Sugimoto, N.; Ozutsumi, K.; Okabe, T.
Infect. Immun. 57, 3588-3593, 1989
A;Title: Isolation, purification, and characterization of fragment B, the NH-2-termina
A;Reference number: A60759; MUID:90035436; PMID:2478476
A;Accession: A60759
A;Molecule type: protein
A;Residues: 461-475 <MAT>
R;Demotz, S.; Lanzavecchia, L.; Eisel, U.; Niemann, H.; Widmann, C.; Corradin, G.
J. Immunol. 142, 394-402, 1989
A;Title: Delineation of several DR-restricted tetanus toxin T cell epitopes.
A;Reference number: JS0098; MUID:89093918; PMID:2463305
A;Contents: annotation; epitope region
R;Schiavo, G.; Benfenati, F.; Poulain, B.; Rossetto, O.; de Laureto, P.P.; DasGupta, B
Nature 359, 832-835, 1992
A;Title: Tetanus and botulinum-B neurotoxins block neurotransmitter release by proteol
A;Reference number: S27125; MUID:93063293; PMID:1331807
A;Contents: annotation

R;de Filippis, V.; Vangelista, L.; Schiavo, G.; Tonello, F.; Montecucco, C.
Eur. J. Biochem. 229, 61-69, 1995
A;Title: Structural studies on the zinc-endopeptidase light chain of tetanus neurotoxin.
A;Reference number: S69348; MUID:95262688; PMID:7744050
A;Accession: S69348
A;Molecule type: protein
A;Residues: 2-31 <DEF>
C;Comment: The source of this protein was an extrachromosomal plasmid.
C;Comment: The precursor is cleaved by endogenous proteinase activity to form light (fra
dual chains are not toxic when separated). The amino end of the heavy chain (fragment B)
C;Comment: Fragment B forms ion channels in a lipid bilayer. Fragment C binds to ganglio
C;Comment: This potent neurotoxin binds to peripheral neuronal synapses, is internalized
presynaptic neurons. It inhibits neurotransmitter release by proteolytic cleavage of sy
C;Function:
A;Description: blocks neuroexocytosis via hydrolysis of a Gln-Phe peptide bond in synapt
C;Superfamily: tetanus toxin
C;Keywords: hydrolase; metalloproteinase; neurotoxin; transmembrane protein; zinc
F;2-457/Product: tentoxylysin light chain (fragment A) #status predicted <TTL>
F;461-1315/Product: tentoxylysin heavy chain (fragment B.C) #status experimental <TTH>
F;461-864/Domain: channel forming (fragment B) #status predicted <TXB>
F;865-1315/Domain: ganglioside binding (fragment C) #status predicted <TXC>
F;233,237/Binding site: zinc (His) #status predicted
F;234/Active site: Glu #status predicted

Query Match 46.9%; Score 76; DB 1; Length 1315;
Best Local Similarity 93.8%; Pred. No. 0.015;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LQYIKANSKFIGITEL 31
:|||||||||||||||
Db 829 MQYIKANSKFIGITEL 844

RESULT 2
RHPGG
gonadoliberin - pig
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 18-Mar-1997
C;Accession: A01411
R;Baba, Y.; Matsuo, H.; Schally, A.V.
Biochem. Biophys. Res. Commun. 44, 459-463, 1971
A;Title: Structure of the porcine LH- and FSH-releasing hormone. II. Confirmation of the
A;Reference number: A90172; MUID:72114303; PMID:4946067
A;Accession: A01411
A;Molecule type: protein
A;Residues: 1-10 <BAB>
R;Matsuo, H.; Arimura, A.; Nair, R.M.G.; Schally, A.V.
Biochem. Biophys. Res. Commun. 45, 822-827, 1971
A;Title: Synthesis of the porcine LH- and FSH-releasing hormone by the solid-phase metho
A;Reference number: A90176; MUID:72065376; PMID:4942726
A;Contents: annotation; synthesis
A;Note: the synthetic and natural hormones have the same physicochemical and biological
R;Baba, Y.; Arimura, A.; Schally, A.V.
Biochem. Biophys. Res. Commun. 45, 483-487, 1971
A;Title: On the tryptophan residue in porcine LH and FSH-releasing hormone.
A;Reference number: A90175; MUID:72117544; PMID:4946275
A;Contents: annotation
A;Note: Trp-3 appears to be essential for biological activity
C;Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and fo
C;Superfamily: gonadoliberin
C;Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 35.8%; Score 58; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.027;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HWSYGLRPG 10
|||||||||
Db 2 HWSYGLRPG 10

RESULT 3
RHSHG
gonadoliberin - sheep
C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Mar-1997
C;Accession: A93780; A01411
R;Burgus, R.; Butcher, M.; Amoss, M.; Ling, N.; Monahan, M.; Rivier, J.; Fellows, R.; B
Proc. Natl. Acad. Sci. U.S.A. 69, 278-282, 1972
A;Title: Primary structure of the ovine hypothalamic luteinizing hormone-releasing fact
A;Reference number: A93780; MUID:72094314; PMID:4550508
A;Accession: A93780
A;Molecule type: protein
A;Residues: 1-10 <BUR>
A;Note: the natural and synthetic hormones have the same biological activity
C;Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and f
C;Superfamily: gonadoliberin
C;Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 35.8%; Score 58; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.027;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HWSYGLRPG 10
|||||||||
Db 2 HWSYGLRPG 10

RESULT 4
I78541
gonadoliberin precursor - rhesus macaque (fragment)
N;Alternate names: luteinizing hormone releasing hormone
C;Species: Macaca mulatta (rhesus macaque)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 16-Jul-1999
C;Accession: I78541
R;Ma, Y.J.; Costa, M.E.; Ojeda, S.R.
Neuroendocrinology 60, 346-359, 1994
A;Title: Developmental expression of the genes encoding transforming growth factor alph
A;Reference number: I58134; MUID:95124501; PMID:7545971
A;Accession: I78541
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-67 <RES>
A;Cross-references: GB:S75918; NID:g912831; PIDN:AAB33096.1; PID:g912832
C;Superfamily: gonadoliberin

Query Match 35.8%; Score 58; DB 2; Length 67;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HWSYGLRPG 10
|||||||||
Db 7 HWSYGLRPG 15

RESULT 5
I51423
gonadoliberin precursor - African clawed frog
N;Alternate names: luteinizing hormone releasing hormone
C;Species: Xenopus laevis (African clawed frog)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 16-Jul-1999
C;Accession: I51423
R;Hayes, W.P.; Wray, S.; Battey, J.F.
Endocrinology 134, 1835-1845, 1994
A;Title: The frog GnRH-I gene has a mammalian-like expression pattern and conserved dom
A;Reference number: I51423; MUID:94185563; PMID:8137750
A;Accession: I51423
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-89 <HAY>
A;Cross-references: GB:L28040; NID:g496291; PIDN:AAA49728.1; PID:g496292
C;Genetics:

A;Gene: GnRH-I
C;Superfamily: gonadoliberin

Query Match 35.8%; Score 58; DB 2; Length 89;
Best Local Similarity 100.0%; Pred. No. 0.31;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HWSYGLRPG 10
||||||||||
Db 25 HWSYGLRPG 33

RESULT 6
RHMSG
gonadoliberin precursor - mouse
N;Alternate names: gonadotropin-releasing hormone (GnRH); luteinizing hormone releasing
N;Contains: gonadoliberin; gonadoliberin-associated protein (GAP)
C;Species: Mus musculus (house mouse)
C;Date: 31-Dec-1993 #sequence_revision 18-Mar-1997 #text_change 18-Jun-1999
C;Accession: A47578
R;Mason, A.J.; Hayflick, J.S.; Zoeller, R.T.; Young III, W.S.; Phillips, H.S.; Nikolics,
Science 234, 1366-1371, 1986
A;Title: A deletion truncating the gonadotropin-releasing hormone gene is responsible fo
A;Reference number: A47578; MUID:87069928; PMID:3024317
A;Accession: A47578
A;Molecule type: DNA
A;Residues: 1-90 <MAS>
A;Cross-references: EMBL:M14872; NID:g193576; PIDN:AAA37717.1; PID:g387175
C;Genetics:
A;Introns: 45/3; 77/3
C;Function:
A;Description: gonadoliberin stimulates pituitary secretion of lutropin and follitropin
A;Note: gonadoliberin-associated protein may have prolactin release inhibiting activity
C;Superfamily: gonadoliberin
C;Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F;1-23/Domain: signal sequence #status predicted <SIG>
F;22-31/Product: gonadoliberin #status predicted <GLB>
F;35-90/Product: gonadoliberin-associated protein #status predicted <GAP>
F;22/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
F;31/Modified site: amidated carboxyl end (Gly) (amide in mature form from following gly

Query Match 35.8%; Score 58; DB 1; Length 90;
Best Local Similarity 100.0%; Pred. No. 0.31;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HWSYGLRPG 10
||||||||||
Db 23 HWSYGLRPG 31

RESULT 7
RHHUG
gonadoliberin precursor [validated] - human
N;Alternate names: gonadotropin releasing hormone (GnRH); luteinizing hormone releasing
N;Contains: gonadoliberin-associated protein (GAP); progonadoliberin
C;Species: Homo sapiens (man)
C;Date: 17-Mar-1987 #sequence_revision 21-Jul-1995 #text_change 08-Dec-2000
C;Accession: S05308; A26173; A93342; A90108; A01410; S45718
R;Hayflick, J.S.; Adelman, J.P.; Seeburg, P.H.
Nucleic Acids Res. 17, 6403-6404, 1989
A;Title: The complete nucleotide sequence of the human gonadotropin-releasing hormone ge
A;Reference number: S05308; MUID:89366682; PMID:2671939
A;Accession: S05308
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-92 <HAY>
A;Cross-references: EMBL:X15215; NID:g31955; PIDN:CAA33285.1; PID:g31956
R;Adelman, J.P.; Mason, A.J.; Hayflick, J.S.; Seeburg, P.H.
Proc. Natl. Acad. Sci. U.S.A. 83, 179-183, 1986
A;Title: Isolation of the gene and hypothalamic cDNA for the common precursor of gonadot
A;Reference number: A94090; MUID:86094338; PMID:2867548
A;Accession: A26173
A;Molecule type: mRNA
A;Residues: 1-92 <ADE>
A;Cross-references: GB:M12578; NID:g183418; PIDN:AAA35916.1; PID:g386749
A;Experimental source: hypothalamus
R;Seeburg, P.H.; Adelman, J.P.
Nature 311, 666-668, 1984
A;Title: Characterization of cDNA for precursor of human luteinizing hormone releasing
A;Reference number: A93342; MUID:85012739; PMID:6090951
A;Accession: A93342
A;Molecule type: mRNA
A;Residues: 1-15,'S',17-92 <SEE>
A;Cross-references: GB:X01059; NID:g34356; PIDN:CAA25526.1; PID:g34357
A;Experimental source: placenta
R;Tan, L.; Rousseau, P.
Biochem. Biophys. Res. Commun. 109, 1061-1071, 1982
A;Title: The chemical identity of the immunoreactive LHRH-like peptide biosynthesized
A;Reference number: A90108; MUID:83126573; PMID:6760865
A;Accession: A90108
A;Molecule type: protein
A;Residues: 24-33 <TAN>
A;Experimental source: placental trophoblasts
R;Leibovitz, D.; Koch, Y.; Pitzer, F.; Fridkin, M.; Dantes, A.; Baumeister, W.; Amster
FEBS Lett. 346, 203-206, 1994
A;Title: Sequential degradation of the neuropeptide gonadotropin-releasing hormone by
A;Reference number: S45718; MUID:94283597; PMID:8013634
A;Contents: annotation; degradation pathway of synthetic hormone
C;Genetics:
A;Gene: GDB:GNRH; LHRH; GRH
A;Cross-references: GDB:133746; OMIM:227200; OMIM:152760
A;Map position: 8p21-8p11.2
A;Introns: 47/3; 79/3
C;Function:
A;Description: gonadoliberin stimulates pituitary secretion of lutropin and follitropi
A;Note: gonadoliberin-associated protein may have prolactin release inhibiting activiy
C;Superfamily: gonadoliberin
C;Keywords: amidated carboxyl end; hormone; hypothalamus; placenta; pyroglutamic acid
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-92/Product: progonadoliberin #status predicted <PGN>
F;24-33/Product: gonadoliberin #status experimental <MAT>
F;37-92/Product: gonadoliberin-associated protein #status predicted <GAP>
F;24/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experim
F;33/Modified site: amidated carboxyl end (Gly) (amide in mature form from following g

Query Match 35.8%; Score 58; DB 1; Length 92;
Best Local Similarity 100.0%; Pred. No. 0.32;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HWSYGLRPG 10
||||||||||
Db 25 HWSYGLRPG 33

RESULT 8
RHRTG
gonadoliberin precursor - rat
N;Alternate names: gonadoliberin-associated protein (GAP); gonadotropin releasing horm
N;Contains: gonadoliberin; prolactin release-inhibiting factor
C;Species: Rattus norvegicus (Norway rat)
C;Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 18-Jun-1999
C;Accession: A40147; B26173; A48410
R;Bond, C.T.; Hayflick, J.S.; Seeburg, P.H.; Adelman, J.P.
Mol. Endocrinol. 3, 1257-1262, 1989
A;Title: The rat gonadotropin-releasing hormone: SH locus: structure and hypothalamic
A;Reference number: A40147; MUID:89384661; PMID:2476669
A;Accession: A40147
A;Molecule type: DNA
A;Residues: 1-92 <BON>
A;Cross-references: GB:M31670; NID:g204447; PIDN:AAA41264.1; PID:g204448
R;Adelman, J.P.; Mason, A.J.; Hayflick, J.S.; Seeburg, P.H.
Proc. Natl. Acad. Sci. U.S.A. 83, 179-183, 1986
A;Title: Isolation of the gene and hypothalamic cDNA for the common precursor of gonad
A;Reference number: A94090; MUID:86094338; PMID:2867548
A;Accession: B26173
A;Molecule type: mRNA

```
A;Residues: 1-92 <ADE>
A;Cross-references: GB:M12579; NID:g204445; PIDN:AAA41263.1; PID:g204446
R;Maier, C.C.; Marchetti, B.; LeBoeuf, R.D.; Blalock, J.E.
Cell. Mol. Neurobiol. 12, 447-454, 1992
A;Title: Thymocytes express a mRNA that is identical to hypothalamic luteinizing hormone
A;Reference number: A48410; MUID:93105480; PMID:1468115
A;Accession: A48410
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-92 <MAI>
A;Cross-references: GB:S50870; NID:g262059; PIDN:AAB24572.1; PID:g262060
A;Experimental source: thymus
A;Note: sequence extracted from NCBI backbone (NCBIN:121082, NCBIP:121083)
C;Genetics:
A;Introns: 47/3; 79/3
C;Function:
A;Description: stimulates pituitary secretion of lutropin and follitropin
A;Note: gonadoliberin-associated protein may have prolactin release inhibiting activiy
C;Superfamily: gonadoliberin
C;Keywords: amidated carboxyl end; hormone; hypothalamus; placenta; pyroglutamic acid; r
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-92/Product: progonadoliberin #status predicted <PGN>
F;24-33/Product: gonadoliberin #status predicted <GLN>
F;37-92/Product: prolactin release-inhibiting factor #status predicted <PIF>
F;24/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
F;33/Modified site: amidated carboxyl end (Gly) (amide in mature form from following gly

  Query Match             35.8%;  Score 58;  DB 1;  Length 92;
  Best Local Similarity   100.0%;  Pred. No. 0.32;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 HWSYGLRPG 10
              |||||||||
Db         25 HWSYGLRPG 33


RESULT 9
T52510
hypothetical protein B2J23.60 [imported] - Neurospora crassa
C;Species: Neurospora crassa
C;Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 20-Oct-2000
C;Accession: T52510
R;Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, September 2000
A;Reference number: Z26053
A;Accession: T52510
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-719 <SCH>
A;Cross-references: EMBL:AL442164; GSPDB:GN00116; NCSP:B2J23.60
A;Experimental source: BAC clone B2J23; strain OR74A
C;Genetics:
A;Gene: NCSP:B2J23.60
A;Map position: 6
A;Introns: 349/1; 601/1

  Query Match             34.6%;  Score 56;  DB 2;  Length 719;
  Best Local Similarity   45.0%;  Pred. No. 6.3;
  Matches    9;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          3 WSYGLRPGSSGPSLQYIKAN 22
              |||  ||||:|  : :: |:
Db        557 WSYSGRPGSAGGLMSFVSAS 576


RESULT 10
S31029
gene 84 protein - Mycobacterium phage L5
C;Species: Mycobacterium phage L5
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 08-Oct-1999
C;Accession: S31029
R;Donnelly-Wu, M.K.; Jacobs Jr., W.R.; Hatfull, G.F.
Mol. Microbiol. 7, 407-417, 1993
A;Title: Superinfection immunity of mycobacteriophage L5: applications for genetic tran
A;Reference number: S30949; MUID:93211283; PMID:8459767
A;Accession: S31029
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-66 <DON>
A;Cross-references: EMBL:Z18946; NID:g15859; PIDN:CAA79460.1; PID:e59702; PID:g579152
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1992
C;Genetics:
A;Gene: 84
A;Start codon: GTG

  Query Match             34.3%;  Score 55.5;  DB 2;  Length 66;
  Best Local Similarity   43.8%;  Pred. No. 0.51;
  Matches   14;  Conservative    2;  Mismatches    7;  Indels    9;  Gaps    2;

Qy          5 YGL-----RPGSSGPSLQYIKANSKFIGITEL 31
              ||      || ||    ||| | ||:|  |:
Db         36 YGFEVDWYEPGESG----YIKRNGKFVGTWEV 63


RESULT 11
G48677
Ig heavy chain V-D-J region (419.1) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 19-May-1994 #sequence_revision 19-May-1994 #text_change 17-Mar-1999
C;Accession: G48677
R;Tassignon, J.; Brait, M.; Jamila, I.; Urbain, J.; Gottlieb, P.; Brown, A.; Hasemann,
Proc. Natl. Acad. Sci. U.S.A. 90, 9508-9512, 1993
A;Title: Molecular characterization of monoclonal CRI-A-positive anti-arsonate antibodi
A;Reference number: A48677; MUID:94022404; PMID:8415731
A;Accession: G48677
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-123 <TAS>
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin
F;15-98/Domain: immunoglobulin homology <IMM>

  Query Match             34.0%;  Score 55;  DB 2;  Length 123;
  Best Local Similarity   40.0%;  Pred. No. 1.2;
  Matches   16;  Conservative    1;  Mismatches   11;  Indels   12;  Gaps    2;

Qy          4 SYGL-----RPGSS-------GPSLQYIKANSKFIGITEL 31
              |||:     |||          |  ||| | || | | ||
Db         31 SYGVNWVKQRPGQGLEWIGYINPGNDYIKYNEKFKGTTTL 70


RESULT 12
RHAQ1
gonadoliberin I - American alligator
N;Alternate names: gonadotropin-releasing hormone I
C;Species: Alligator mississippiensis (American alligator)
C;Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 18-Mar-1997
C;Accession: A60066
R;Lovejoy, D.A.; Fischer, W.H.; Parker, D.B.; McRory, J.E.; Park, M.; Lance, V.; Swansc
Regul. Pept. 33, 105-116, 1991
A;Title: Primary structure of two forms of gonadotropin-releasing hormone from brains c
A;Reference number: A60066; MUID:91352338; PMID:1882082
A;Accession: A60066
A;Molecule type: protein
A;Residues: 1-10 <LOV>
C;Superfamily: gonadoliberin
C;Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;10/Modified site: amidated carboxyl end (Gly) #status experimental

  Query Match             33.3%;  Score 54;  DB 1;  Length 10;
  Best Local Similarity   88.9%;  Pred. No. 0.1;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 HWSYGLRPG 10
              ||||||:||
```

```
Db          2 HWSYGLQPG 10


RESULT 13
I50644
gonadoliberin I precursor - chicken
N;Alternate names: gonadotropin-releasing hormone I
C;Species: Gallus gallus (chicken)
C;Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 16-Jul-1999
C;Accession: I50644; S33507
R;Dunn, I.C.; Chen, Y.; Hook, C.; Sharp, P.J.; Sang, H.M.
J. Mol. Endocrinol. 11, 19-29, 1993
A;Title: Characterization of the chicken preprogonadotrophin-releasing hormone-I gene.
A;Reference number: I50644; MUID:94059355; PMID:7902095
A;Accession: I50644
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-92 <DU2>
A;Cross-references: EMBL:X69491; NID:g496326; PIDN:CAA49246.1; PID:g311612
C;Genetics:
A;Introns: 47/3; 79/3
C;Superfamily: gonadoliberin

  Query Match            33.3%;  Score 54;  DB 2;  Length 92;
  Best Local Similarity   88.9%;  Pred. No. 1.2;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 HWSYGLRPG 10
              ||||||:||
Db         25 HWSYGLQPG 33


RESULT 14
E95361
probable muconate cycloisomerase (EC 5.5.1.1) [imported] - Sinorhizobium meliloti (strai
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C;Accession: E95361
R;Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows
.; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A;Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium melilot
A;Reference number: A95262; MUID:21396509; PMID:11481432
A;Accession: E95361
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-374 <KUR>
A;Cross-references: GB:AE006469; PIDN:AAK65455.1; PID:g14523923; GSPDB:GN00165
A;Experimental source: strain 1021, megaplasmid pSymA
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
 L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: SMa1461
A;Genome: plasmid
C;Keywords: intramolecular lyase; isomerase

  Query Match            33.0%;  Score 53.5;  DB 2;  Length 374;
  Best Local Similarity   44.8%;  Pred. No. 7;
  Matches   13;  Conservative    5;  Mismatches   10;  Indels    1;  Gaps    1;

Qy          2 HWSYGLRPGSSGPSLQYIKANSKFIGITE 30
              |||||:|  |   :|  |:|:   :|| |
Db         21 HWSYGIRE-SFAVNLIEIEADDGTVGIGE 48


RESULT 15
I50739
gonadotropin-releasing hormone - Cichlid (Haplochromis burtoni)
C;Species: Haplochromis burtoni
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
C;Accession: I50739
R;White, S.A.; Kasten, T.L.; Bond, C.T.; Adelman, J.P.; Fernald, R.D.
Proc. Natl. Acad. Sci. U.S.A. 92, 8363-8367, 1995
A;Title: Three gonadotropin-releasing hormone genes in one organism suggest novel role
A;Reference number: I50739; MUID:95396797; PMID:7667296
A;Accession: I50739
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-98 <WHI>
A;Cross-references: EMBL:U31865; NID:g905398; PIDN:AAC59691.1; PID:g905399
C;Superfamily: gonadoliberin

  Query Match            32.1%;  Score 52;  DB 2;  Length 98;
  Best Local Similarity   88.9%;  Pred. No. 2.6;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 HWSYGLRPG 10
              ||||||:||
Db         24 HWSYGLSPG 32


Search completed: March 10, 2004, 09:16:50
Job time : 9.70863 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:54 ; Search time 14.3249 Seconds
(without alignments)
133.345 Million cell updates/sec

Title: US-09-848-834A-14
Perfect score: 200
Sequence: 1 XHWSYGLRPGSSGPSLFNNFTVSFWLRVPKVSASHLE 37

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	116	58.0	188	4	US-09-396-937-14	Sequence 14, Appl
2	114	57.0	31	5	PCT-US93-11703-64	Sequence 64, Appl
3	114	57.0	452	1	US-07-618-312A-2	Sequence 2, Appli
4	114	57.0	452	1	US-07-618-312A-4	Sequence 4, Appli
5	114	57.0	452	1	US-08-110-786A-8	Sequence 8, Appli
6	114	57.0	452	1	US-08-280-228-2	Sequence 2, Appli
7	114	57.0	452	1	US-08-280-228-4	Sequence 4, Appli
8	114	57.0	618	1	US-08-668-381A-5	Sequence 5, Appli
9	114	57.0	853	4	US-08-913-880C-17	Sequence 17, Appl
10	114	57.0	858	4	US-08-913-880C-16	Sequence 16, Appl
11	114	57.0	860	4	US-08-913-880C-15	Sequence 15, Appl
12	114	57.0	862	4	US-08-913-880C-14	Sequence 14, Appl
13	114	57.0	865	4	US-08-913-880C-13	Sequence 13, Appl
14	114	57.0	866	4	US-08-913-880C-12	Sequence 12, Appl
15	114	57.0	874	4	US-08-913-880C-11	Sequence 11, Appl
16	114	57.0	875	4	US-08-913-880C-10	Sequence 10, Appl
17	114	57.0	1315	4	US-08-913-880C-1	Sequence 1, Appli
18	112	56.0	21	1	US-07-610-525-1	Sequence 1, Appli
19	112	56.0	21	2	US-08-661-052-12	Sequence 12, Appl
20	112	56.0	21	2	US-08-460-502-8	Sequence 8, Appli
21	112	56.0	21	2	US-08-724-774B-5	Sequence 5, Appli
22	112	56.0	21	3	US-09-089-595-5	Sequence 5, Appli
23	112	56.0	21	3	US-09-382-855-5	Sequence 5, Appli
24	112	56.0	21	3	US-09-183-714B-5	Sequence 5, Appli
25	112	56.0	21	3	US-09-188-082-12	Sequence 12, Appl
26	112	56.0	21	3	US-09-171-969-10	Sequence 10, Appl
27	112	56.0	21	4	US-09-364-088-12	Sequence 12, Appl
28	112	56.0	21	4	US-09-642-281-5	Sequence 5, Appli
29	112	56.0	21	4	US-09-102-716-12	Sequence 12, Appl
30	112	56.0	21	4	US-08-432-483A-3	Sequence 3, Appli
31	112	56.0	21	4	US-09-148-711A-8	Sequence 8, Appli
32	112	56.0	21	4	US-09-589-717-5	Sequence 5, Appli
33	112	56.0	21	4	US-08-945-289-3	Sequence 3, Appli
34	112	56.0	21	4	US-09-396-937-35	Sequence 35, Appl
35	112	56.0	21	4	US-09-405-986A-2	Sequence 2, Appli
36	112	56.0	21	5	PCT-US93-11703-66	Sequence 66, Appl
37	112	56.0	32	1	US-08-446-692-14	Sequence 14, Appl
38	112	56.0	32	2	US-08-488-351A-14	Sequence 14, Appl
39	112	56.0	173	4	US-09-396-937-20	Sequence 20, Appl
40	107	53.5	22	1	US-08-446-692-5	Sequence 5, Appli
41	107	53.5	22	2	US-08-488-351A-5	Sequence 5, Appli
42	107	53.5	22	3	US-09-100-409A-41	Sequence 41, Appl
43	107	53.5	22	5	PCT-US95-13841-8	Sequence 8, Appli
44	100	50.0	19	1	US-07-610-525-2	Sequence 2, Appli
45	94.5	47.2	20	2	US-08-319-704-11	Sequence 11, Appl

ALIGNMENTS

```
RESULT 1
US-09-396-937-14
; Sequence 14, Application US/09396937
; Patent No. 6645500
; GENERAL INFORMATION:
;  APPLICANT: M&E Biotech A/S
;  APPLICANT: HALKIER, Torben
;  APPLICANT: HAANING, Jesper
;  TITLE OF INVENTION: Method for Down-Regulating Osteoprotegerin Ligand
;  TITLE OF INVENTION: Activity
;  FILE REFERENCE: 22021 PC 1
;  CURRENT APPLICATION NUMBER: US/09/396,937
;  CURRENT FILING DATE: 1999-09-15
;  NUMBER OF SEQ ID NOS: 35
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
;   LENGTH: 188
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Fusion of
;   OTHER INFORMATION: murine OPGL, residues 158-316 modified by
;   OTHER INFORMATION: introduction of tetanus toxoid P30 epitope, and
;   OTHER INFORMATION: His tag
US-09-396-937-14

  Query Match           58.0%;  Score 116;  DB 4;  Length 188;
  Best Local Similarity  85.2%;  Pred. No. 1.3e-09;
  Matches   23;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         11 SSGPSLFNNFTVSFWLRVPKVSASHLE 37
              ||   :||||||||||||||||||||||
Db        107 SSHNLMFNNFTVSFWLRVPKVSASHLE 133


RESULT 2
PCT-US93-11703-64
; Sequence 64, Application PC/TUS9311703
;  GENERAL INFORMATION:
;    APPLICANT:  Chiron Mimotopes Pty. Ltd.
;    TITLE OF INVENTION:  T-Cell Epitopes
;    NUMBER OF SEQUENCES:  75
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Grant D. Green
;      STREET:  4560 Horton St.
;      CITY:  Emeryville
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94608
;    COMPUTER READABLE FORM:
```

```
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30B
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US93/11703
;       FILING DATE:  28-DEC-1993
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/984,852
;       FILING DATE:  02-DEC-1992
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Green, Grant D.
;       REGISTRATION NUMBER:  31,259
;       REFERENCE/DOCKET NUMBER:  0222.101
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  510-601-2706
;       TELEFAX:  510-655-3542
;   INFORMATION FOR SEQ ID NO:  64:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  31 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
PCT-US93-11703-64

  Query Match             57.0%;  Score 114;  DB 5;  Length 31;
  Best Local Similarity   95.5%;  Pred. No. 2.9e-10;
  Matches   21;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         16 LFNNFTVSFWLRVPKVSASHLE 37
              :||||||||||||||||||||||
Db          6 MFNNFTVSFWLRVPKVSASHLE 27


RESULT 3
US-07-618-312A-2
; Sequence 2, Application US/07618312A
; Patent No. 5389540
;  GENERAL INFORMATION:
;    APPLICANT:  Makoff Dr, Andrew J
;    APPLICANT:  Romanos Dr, Michael A
;    APPLICANT:  Clare Dr, Jeffrey J
;    APPLICANT:  Fairweather Dr, Neil F
;    TITLE OF INVENTION:  VACCINES
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  14th Floor
;      STREET:  2200 Clarendon Boulevard,
;      CITY:  Arlington,
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22201
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/618,312A
;      FILING DATE:  19910516
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 8926832.0
;      FILING DATE:  28-NOV-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9006097.1
;      FILING DATE:  17-MAR-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Crawford Mr, Arthur R
;      REGISTRATION NUMBER:  25,327
;      REFERENCE/DOCKET NUMBER:  510-51
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  0101 703 8750400
;      TELEFAX:  0101 703 5253468
;      TELEX:  200797 NIXN UR
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  452 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-07-618-312A-2

  Query Match             57.0%;  Score 114;  DB 1;  Length 452;
  Best Local Similarity   95.5%;  Pred. No. 7.6e-09;
  Matches   21;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         16 LFNNFTVSFWLRVPKVSASHLE 37
              :||||||||||||||||||||||
Db         83 MFNNFTVSFWLRVPKVSASHLE 104


RESULT 4
US-07-618-312A-4
; Sequence 4, Application US/07618312A
; Patent No. 5389540
;  GENERAL INFORMATION:
;    APPLICANT:  Makoff Dr, Andrew J
;    APPLICANT:  Romanos Dr, Michael A
;    APPLICANT:  Clare Dr, Jeffrey J
;    APPLICANT:  Fairweather Dr, Neil F
;    TITLE OF INVENTION:  VACCINES
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  14th Floor
;      STREET:  2200 Clarendon Boulevard,
;      CITY:  Arlington,
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22201
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/618,312A
;      FILING DATE:  19910516
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 8926832.0
;      FILING DATE:  28-NOV-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9006097.1
;      FILING DATE:  17-MAR-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Crawford Mr, Arthur R
;      REGISTRATION NUMBER:  25,327
;      REFERENCE/DOCKET NUMBER:  510-51
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  0101 703 8750400
;      TELEFAX:  0101 703 5253468
;      TELEX:  200797 NIXN UR
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  452 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-07-618-312A-4

  Query Match             57.0%;  Score 114;  DB 1;  Length 452;
  Best Local Similarity   95.5%;  Pred. No. 7.6e-09;
  Matches   21;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          16 LFNNFTVSFWLRVPKVSASHLE 37
               :||||||||||||||||||||||
Db          83 MFNNFTVSFWLRVPKVSASHLE 104


RESULT 5
US-08-110-786A-8
; Sequence 8, Application US/08110786A
; Patent No. 5443966
;  GENERAL INFORMATION:
;    APPLICANT:  FAIRWEATHER, Neil Fraser
;    APPLICANT:  MAKOFF, Andrew Joseph
;    TITLE OF INVENTION:  Expression of tetanus toxin fragment C
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Nixon & Vanderhye P.C.
;      STREET:  1100 No. 5443966th Glebe Road
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22201-4714
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/110,786A
;      FILING DATE:  23-AUG-1993 1991
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/777,337
;      FILING DATE:  29-NOV-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB90/00943
;      FILING DATE:  20-JUN-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 8914122.0
;      FILING DATE:  20 June 1989
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mary J. Wilson
;      REGISTRATION NUMBER:  32,955
;      REFERENCE/DOCKET NUMBER:  117-134
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 816-4000
;      TELEFAX:  (703) 816-4100
;      TELEX:  200797 NIXN UR
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  452 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-110-786A-8

  Query Match             57.0%;  Score 114;  DB 1;  Length 452;
  Best Local Similarity   95.5%;  Pred. No. 7.6e-09;
  Matches   21;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          16 LFNNFTVSFWLRVPKVSASHLE 37
               :||||||||||||||||||||||
Db          83 MFNNFTVSFWLRVPKVSASHLE 104


RESULT 6
US-08-280-228-2
; Sequence 2, Application US/08280228
; Patent No. 5571694
;  GENERAL INFORMATION:
;    APPLICANT:  Makoff Dr, Andrew J
;    APPLICANT:  Romanos Dr, Michael A
;    APPLICANT:  Clare Dr, Jeffrey J
;    APPLICANT:  Fairweather Dr, Neil F
;    TITLE OF INVENTION:  VACCINES
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NIXON & VANDERHYE P.C.
;      STREET:  1100 No. 5571694th Glebe Road
;      CITY:  Arlington,
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22201-4714
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/280,228
;      FILING DATE:  25-JUL-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/618,312
;      FILING DATE:  27-NOV-1990
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 8926832.0
;      FILING DATE:  28-NOV-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9006097.1
;      FILING DATE:  17-MAR-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Wilson, Mary J.
;      REGISTRATION NUMBER:  32,955
;      REFERENCE/DOCKET NUMBER:  117-163
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 816-4000
;      TELEFAX:  (703) 816-4100
;      TELEX:  200797 NIXN UR
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  452 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-280-228-2

  Query Match             57.0%;  Score 114;  DB 1;  Length 452;
  Best Local Similarity   95.5%;  Pred. No. 7.6e-09;
  Matches   21;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          16 LFNNFTVSFWLRVPKVSASHLE 37
               :||||||||||||||||||||||
Db          83 MFNNFTVSFWLRVPKVSASHLE 104


RESULT 7
US-08-280-228-4
; Sequence 4, Application US/08280228
; Patent No. 5571694
;  GENERAL INFORMATION:
;    APPLICANT:  Makoff Dr, Andrew J
;    APPLICANT:  Romanos Dr, Michael A
;    APPLICANT:  Clare Dr, Jeffrey J
;    APPLICANT:  Fairweather Dr, Neil F
;    TITLE OF INVENTION:  VACCINES
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NIXON & VANDERHYE P.C.
;      STREET:  1100 No. 5571694th Glebe Road
;      CITY:  Arlington,
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22201-4714
;    COMPUTER READABLE FORM:
```

```
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/280,228
;       FILING DATE:  25-JUL-1994
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/618,312
;       FILING DATE:  27-NOV-1990
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  GB 8926832.0
;       FILING DATE:  28-NOV-1989
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  GB 9006097.1
;       FILING DATE:  17-MAR-1990
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Wilson, Mary J.
;       REGISTRATION NUMBER:  32,955
;       REFERENCE/DOCKET NUMBER:  117-163
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (703) 816-4000
;       TELEFAX:  (703) 816-4100
;       TELEX:  200797 NIXN UR
;   INFORMATION FOR SEQ ID NO:  4:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  452 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-280-228-4

  Query Match             57.0%;  Score 114;  DB 1;  Length 452;
  Best Local Similarity   95.5%;  Pred. No. 7.6e-09;
  Matches   21;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         16 LFNNFTVSFWLRVPKVSASHLE 37
              :||||||||||||||||||||||
Db         83 MFNNFTVSFWLRVPKVSASHLE 104


RESULT 8
US-08-668-381A-5
; Sequence 5, Application US/08668381A
; Patent No. 5780024
;  GENERAL INFORMATION:
;     APPLICANT:  Brown, Robert H.
;     APPLICANT:  Fishman, Paul S.
;     APPLICANT:  Francis, Jonathan W.
;     APPLICANT:  Hosler, Betsy A.
;     TITLE OF INVENTION:  SUPEROXIDE DISMUTASE/TETANUS TOXIN
;     TITLE OF INVENTION:  FRAGMENT C HYBRID PROTEIN
;     NUMBER OF SEQUENCES:  6
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Fish & Richardson P.C.
;       STREET:  225 Franklin Street
;       CITY:  Boston
;       STATE:  MA
;       COUNTRY:  USA
;       ZIP:  02110-2804
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/668,381A
;       FILING DATE:  21-JUN-1996
;       CLASSIFICATION:  514
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  60/000,473
;       FILING DATE:  23-JUN-1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Clark, Paul T.
;       REGISTRATION NUMBER:  30,164
;       REFERENCE/DOCKET NUMBER:  00786/269001
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  617/542-5070
;       TELEFAX:  617/542-8906
;       TELEX:  200154
;   INFORMATION FOR SEQ ID NO:  5:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  618 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-668-381A-5

  Query Match             57.0%;  Score 114;  DB 1;  Length 618;
  Best Local Similarity   95.5%;  Pred. No. 1.1e-08;
  Matches   21;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         16 LFNNFTVSFWLRVPKVSASHLE 37
              :||||||||||||||||||||||
Db        249 MFNNFTVSFWLRVPKVSASHLE 270


RESULT 9
US-08-913-880C-17
; Sequence 17, Application US/08913880C
; Patent No. 6372225
; GENERAL INFORMATION:
;  APPLICANT: MATSUDA, Morihiro
;  TITLE OF INVENTION: TETANUS TOXIN FUNCTIONAL FRAGMENT ANTIGEN AND TETANUS
;  TITLE OF INVENTION:  VACCINE
;  FILE REFERENCE: 216-380P
;  CURRENT APPLICATION NUMBER: US/08/913,880C
;  CURRENT FILING DATE:  1997-11-24
;  NUMBER OF SEQ ID NOS: 17
; SEQ ID NO 17
;   LENGTH: 853
;   TYPE: PRT
;   ORGANISM: Clostridium tetani
;   FEATURE:
;   OTHER INFORMATION: Amino Acids 463 to 1315 of SEQ ID NO: 1
US-08-913-880C-17

  Query Match             57.0%;  Score 114;  DB 4;  Length 853;
  Best Local Similarity   95.5%;  Pred. No. 1.6e-08;
  Matches   21;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         16 LFNNFTVSFWLRVPKVSASHLE 37
              :||||||||||||||||||||||
Db        484 MFNNFTVSFWLRVPKVSASHLE 505


RESULT 10
US-08-913-880C-16
; Sequence 16, Application US/08913880C
; Patent No. 6372225
; GENERAL INFORMATION:
;  APPLICANT: MATSUDA, Morihiro
;  TITLE OF INVENTION: TETANUS TOXIN FUNCTIONAL FRAGMENT ANTIGEN AND TETANUS
;  TITLE OF INVENTION:  VACCINE
;  FILE REFERENCE: 216-380P
;  CURRENT APPLICATION NUMBER: US/08/913,880C
;  CURRENT FILING DATE:  1997-11-24
;  NUMBER OF SEQ ID NOS: 17
; SEQ ID NO 16
;   LENGTH: 858
;   TYPE: PRT
;   ORGANISM: Clostridium tetani
;   FEATURE:
;   OTHER INFORMATION: Amino Acids 458 to 1315 of SEQ ID NO: 1
```

US-08-913-880C-16

Query Match 57.0%; Score 114; DB 4; Length 858;
Best Local Similarity 95.5%; Pred. No. 1.6e-08;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LFNNFTVSFWLRVPKVSASHLE 37
:||||||||||||||||||||||
Db 489 MFNNFTVSFWLRVPKVSASHLE 510

RESULT 11
US-08-913-880C-15
; Sequence 15, Application US/08913880C
; Patent No. 6372225
; GENERAL INFORMATION:
; APPLICANT: MATSUDA, Morihiro
; TITLE OF INVENTION: TETANUS TOXIN FUNCTIONAL FRAGMENT ANTIGEN AND TETANUS
; TITLE OF INVENTION: VACCINE
; FILE REFERENCE: 216-380P
; CURRENT APPLICATION NUMBER: US/08/913,880C
; CURRENT FILING DATE: 1997-11-24
; NUMBER OF SEQ ID NOS: 17
; SEQ ID NO 15
; LENGTH: 860
; TYPE: PRT
; ORGANISM: Clostridium tetani
; FEATURE:
; OTHER INFORMATION: Amino Acids 456 to 1315 of SEQ ID NO: 1
US-08-913-880C-15

Query Match 57.0%; Score 114; DB 4; Length 860;
Best Local Similarity 95.5%; Pred. No. 1.7e-08;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LFNNFTVSFWLRVPKVSASHLE 37
:||||||||||||||||||||||
Db 491 MFNNFTVSFWLRVPKVSASHLE 512

RESULT 12
US-08-913-880C-14
; Sequence 14, Application US/08913880C
; Patent No. 6372225
; GENERAL INFORMATION:
; APPLICANT: MATSUDA, Morihiro
; TITLE OF INVENTION: TETANUS TOXIN FUNCTIONAL FRAGMENT ANTIGEN AND TETANUS
; TITLE OF INVENTION: VACCINE
; FILE REFERENCE: 216-380P
; CURRENT APPLICATION NUMBER: US/08/913,880C
; CURRENT FILING DATE: 1997-11-24
; NUMBER OF SEQ ID NOS: 17
; SEQ ID NO 14
; LENGTH: 862
; TYPE: PRT
; ORGANISM: Clostridium tetani
; FEATURE:
; OTHER INFORMATION: Amino Acids 454 to 1315 of SEQ ID NO: 1
US-08-913-880C-14

Query Match 57.0%; Score 114; DB 4; Length 862;
Best Local Similarity 95.5%; Pred. No. 1.7e-08;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LFNNFTVSFWLRVPKVSASHLE 37
:||||||||||||||||||||||
Db 493 MFNNFTVSFWLRVPKVSASHLE 514

RESULT 13
US-08-913-880C-13
; Sequence 13, Application US/08913880C
; Patent No. 6372225
; GENERAL INFORMATION:
; APPLICANT: MATSUDA, Morihiro
; TITLE OF INVENTION: TETANUS TOXIN FUNCTIONAL FRAGMENT ANTIGEN AND TETANUS
; TITLE OF INVENTION: VACCINE
; FILE REFERENCE: 216-380P
; CURRENT APPLICATION NUMBER: US/08/913,880C
; CURRENT FILING DATE: 1997-11-24
; NUMBER OF SEQ ID NOS: 17
; SEQ ID NO 13
; LENGTH: 865
; TYPE: PRT
; ORGANISM: Clostridium tetani
; FEATURE:
; OTHER INFORMATION: Amino Acids 451 to 1315 of SEQ ID NO: 1
US-08-913-880C-13

Query Match 57.0%; Score 114; DB 4; Length 865;
Best Local Similarity 95.5%; Pred. No. 1.7e-08;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LFNNFTVSFWLRVPKVSASHLE 37
:||||||||||||||||||||||
Db 496 MFNNFTVSFWLRVPKVSASHLE 517

RESULT 14
US-08-913-880C-12
; Sequence 12, Application US/08913880C
; Patent No. 6372225
; GENERAL INFORMATION:
; APPLICANT: MATSUDA, Morihiro
; TITLE OF INVENTION: TETANUS TOXIN FUNCTIONAL FRAGMENT ANTIGEN AND TETANUS
; TITLE OF INVENTION: VACCINE
; FILE REFERENCE: 216-380P
; CURRENT APPLICATION NUMBER: US/08/913,880C
; CURRENT FILING DATE: 1997-11-24
; NUMBER OF SEQ ID NOS: 17
; SEQ ID NO 12
; LENGTH: 866
; TYPE: PRT
; ORGANISM: Clostridium tetani
; FEATURE:
; OTHER INFORMATION: Amino Acids 450 to 1315 of SEQ ID NO: 1
US-08-913-880C-12

Query Match 57.0%; Score 114; DB 4; Length 866;
Best Local Similarity 95.5%; Pred. No. 1.7e-08;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LFNNFTVSFWLRVPKVSASHLE 37
:||||||||||||||||||||||
Db 497 MFNNFTVSFWLRVPKVSASHLE 518

RESULT 15
US-08-913-880C-11
; Sequence 11, Application US/08913880C
; Patent No. 6372225
; GENERAL INFORMATION:
; APPLICANT: MATSUDA, Morihiro
; TITLE OF INVENTION: TETANUS TOXIN FUNCTIONAL FRAGMENT ANTIGEN AND TETANUS
; TITLE OF INVENTION: VACCINE
; FILE REFERENCE: 216-380P
; CURRENT APPLICATION NUMBER: US/08/913,880C
; CURRENT FILING DATE: 1997-11-24
; NUMBER OF SEQ ID NOS: 17
; SEQ ID NO 11
; LENGTH: 874
; TYPE: PRT
; ORGANISM: Clostridium tetani
; FEATURE:
; OTHER INFORMATION: Amino Acids 442 to 1315 of SEQ ID NO: 1
US-08-913-880C-11

Query Match 57.0%; Score 114; DB 4; Length 874;
Best Local Similarity 95.5%; Pred. No. 1.7e-08;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy         16 LFNNFTVSFWLRVPKVSASHLE 37
              :||||||||||||||||||||||
Db        505 MFNNFTVSFWLRVPKVSASHLE 526
```

Search completed: March 10, 2004, 09:28:55
Job time : 14.3249 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 10, 2004, 09:16:59 ; Search time 29.0817 Seconds
(without alignments)
268.645 Million cell updates/sec

Title: US-09-848-834A-14
Perfect score: 200
Sequence: 1 XHWSYGLRPGSSGPSLFNNFTVSFWLRVPKVSASHLE 37

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	199	99.5	37	9	US-09-848-834A-14	Sequence 14, Appl
2	199	99.5	50	9	US-09-848-834A-18	Sequence 18, Appl
3	118.5	59.2	158	14	US-10-297-942-18	Sequence 18, Appl
4	117	58.5	287	14	US-10-295-074-13	Sequence 13, Appl
5	117	58.5	514	14	US-10-295-074-59	Sequence 59, Appl
6	116	58.0	158	14	US-10-297-942-2	Sequence 2, Appli
7	116	58.0	194	14	US-10-295-074-46	Sequence 46, Appl
8	116	58.0	285	14	US-10-295-074-11	Sequence 11, Appl
9	115	57.5	194	14	US-10-295-074-47	Sequence 47, Appl
10	114	57.0	441	15	US-10-452-024-180	Sequence 180, App
11	114	57.0	441	15	US-10-452-024-183	Sequence 183, App
12	114	57.0	444	15	US-10-452-024-185	Sequence 185, App
13	114	57.0	451	15	US-10-452-024-186	Sequence 186, App
14	114	57.0	452	15	US-10-452-024-184	Sequence 184, App
15	114	57.0	463	10	US-09-816-467-2	Sequence 2, Appli
16	114	57.0	469	15	US-10-452-024-182	Sequence 182, App
17	114	57.0	472	15	US-10-452-024-181	Sequence 181, App
18	114	57.0	573	15	US-10-452-024-177	Sequence 177, App
19	114	57.0	605	14	US-10-130-973A-11	Sequence 11, Appl
20	114	57.0	665	14	US-10-130-973A-9	Sequence 9, Appli
21	114	57.0	882	14	US-10-130-973A-3	Sequence 3, Appli
22	114	57.0	907	14	US-10-130-973A-5	Sequence 5, Appli
23	114	57.0	1052	14	US-10-130-973A-17	Sequence 17, Appl
24	114	57.0	1112	14	US-10-130-973A-16	Sequence 16, Appl
25	114	57.0	1310	15	US-10-452-024-149	Sequence 149, App
26	114	57.0	1315	14	US-10-241-596-141	Sequence 141, App
27	114	57.0	1315	15	US-10-452-024-145	Sequence 145, App
28	112	56.0	21	9	US-09-943-548-3	Sequence 3, Appli
29	112	56.0	21	9	US-09-848-834A-4	Sequence 4, Appli
30	112	56.0	21	9	US-09-785-215-6	Sequence 6, Appli
31	112	56.0	21	10	US-09-405-986-2	Sequence 2, Appli
32	112	56.0	21	14	US-10-204-362-6	Sequence 6, Appli
33	112	56.0	21	14	US-10-339-522-3	Sequence 3, Appli
34	112	56.0	21	14	US-10-223-711-8	Sequence 8, Appli
35	112	56.0	21	14	US-10-223-809A-6	Sequence 6, Appli
36	112	56.0	21	14	US-10-261-208-5	Sequence 5, Appli
37	112	56.0	21	14	US-10-295-074-5	Sequence 5, Appli
38	112	56.0	21	15	US-10-372-111-8	Sequence 8, Appli
39	112	56.0	34	9	US-09-848-834A-10	Sequence 10, Appl
40	112	56.0	158	14	US-10-297-942-10	Sequence 10, Appl
41	112	56.0	158	14	US-10-297-942-12	Sequence 12, Appl
42	112	56.0	158	14	US-10-297-942-20	Sequence 20, Appl
43	112	56.0	285	14	US-10-295-074-9	Sequence 9, Appli
44	112	56.0	287	14	US-10-295-074-15	Sequence 15, Appl
45	112	56.0	514	14	US-10-295-074-49	Sequence 49, Appl

ALIGNMENTS

RESULT 1
US-09-848-834A-14
; Sequence 14, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Aphton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
; LENGTH: 37
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the
; OTHER INFORMATION: nRH hormone linked by a spacer to amino acid sequence 947-967 c
; OTHER INFORMATION: the Tetanus toxoid precursor (Tentoxylysin)
; NAME/KEY: MOD_RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(10)
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
; NAME/KEY: PEPTIDE
; LOCATION: (11)..(16)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (17)..(37)
; OTHER INFORMATION: Amino acid sequence 947-967 of the Tetanus toxoid precursor
; OTHER INFORMATION: (Tentoxylysin)
US-09-848-834A-14

Query Match 99.5%; Score 199; DB 9; Length 37;

```
  Best Local Similarity   100.0%;  Pred. No. 4.4e-20;
  Matches   36;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 HWSYGLRPGSSGPSLFNNFTVSFWLRVPKVSASHLE 37
              |||||||||||||||||||||||||||||||||||||
Db          2 HWSYGLRPGSSGPSLFNNFTVSFWLRVPKVSASHLE 37


RESULT 2
US-09-848-834A-18
; Sequence 18, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
;  APPLICANT: Aphton Corporation
;  TITLE OF INVENTION: Chimeric Peptide Immunogens
;  FILE REFERENCE: 1102865-0047
;  CURRENT APPLICATION NUMBER: US/09/848,834A
;  CURRENT FILING DATE:  2001-05-04
;  PRIOR APPLICATION NUMBER: 60/202,328
;  PRIOR FILING DATE: 2000-05-05
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18
;   LENGTH: 50
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of human
;   OTHER INFORMATION: GnRH linked by a spacer to amino acid sequence 947-967 of the Tet
;   OTHER INFORMATION: anus toxoid precursor (Tentoxylysin) protein linked by a spacer t
;   OTHER INFORMATION: o amino acid sequence 2-10 of human GnRH
;   NAME/KEY: MOD_RES
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
;   NAME/KEY: MOD_RES
;   LOCATION: (50)..(50)
;   OTHER INFORMATION: Amidated glycine or glycinamide
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)..(10)
;   OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
;   NAME/KEY: PEPTIDE
;   LOCATION: (11)..(16)
;   OTHER INFORMATION: Spacer peptide
;   NAME/KEY: PEPTIDE
;   LOCATION: (17)..(37)
;   OTHER INFORMATION: Amino acid sequence 947-967 of the Tetanus toxoid precursor (Tent
;   OTHER INFORMATION: oxylysin
;   NAME/KEY: PEPTIDE
;   LOCATION: (38)..(41)
;   OTHER INFORMATION: Spacer peptide
;   NAME/KEY: PEPTIDE
;   LOCATION: (42)..(50)
;   OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone
US-09-848-834A-18

  Query Match             99.5%;  Score 199;  DB 9;  Length 50;
  Best Local Similarity   100.0%;  Pred. No. 6.1e-20;
  Matches   36;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 HWSYGLRPGSSGPSLFNNFTVSFWLRVPKVSASHLE 37
              |||||||||||||||||||||||||||||||||||||
Db          2 HWSYGLRPGSSGPSLFNNFTVSFWLRVPKVSASHLE 37


RESULT 3
US-10-297-942-18
; Sequence 18, Application US/10297942
; Publication No. US20030185816A1
; GENERAL INFORMATION:
;  APPLICANT: Ferring BV
;  TITLE OF INVENTION: Solubilised Protein Vaccines
;  FILE REFERENCE: P68445US0
;  CURRENT APPLICATION NUMBER: US/10/297,942
;  CURRENT FILING DATE:  2003-04-21
;  PRIOR APPLICATION NUMBER: PCT/DK01/00431
;  PRIOR FILING DATE: 2001-10-16
;  PRIOR APPLICATION NUMBER: DK PA 2000 00966
;  PRIOR FILING DATE: 2000-06-21
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
;   LENGTH: 158
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-297-942-18

  Query Match             59.2%;  Score 118.5;  DB 14;  Length 158;
  Best Local Similarity   80.6%;  Pred. No. 2.2e-08;
  Matches   25;  Conservative    1;  Mismatches    4;  Indels    1;  Gaps    1;

Qy          7 LRPGSSGPSLFNNFTVSFWLRVPKVSASHLE 37
              :|  |  || ||||||||||||||||||||||
Db          2 VRSSSRTPS-FNNFTVSFWLRVPKVSASHLE 31


RESULT 4
US-10-295-074-13
; Sequence 13, Application US/10295074
; Publication No. US20030185845A1
; GENERAL INFORMATION:
;  APPLICANT: Pharmexa A/S
;  TITLE OF INVENTION: NOVEL IMMUNOGENIC MIMETICS OF MULTIMER PROTEINS
;  FILE REFERENCE: P1013DK00
;  CURRENT APPLICATION NUMBER: US/10/295,074
;  CURRENT FILING DATE:  2002-11-15
;  NUMBER OF SEQ ID NOS: 60
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
;   LENGTH: 287
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: Two human IL5 monomers joined by diglycine linker and including
;   OTHER INFORMATION: erminally positioned P30 and P2 epitopes
US-10-295-074-13

  Query Match             58.5%;  Score 117;  DB 14;  Length 287;
  Best Local Similarity   91.7%;  Pred. No. 6.9e-08;
  Matches   22;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         14 PSLFNNFTVSFWLRVPKVSASHLE 37
              |: ||||||||||||||||||||||
Db         21 PTEFNNFTVSFWLRVPKVSASHLE 44


RESULT 5
US-10-295-074-59
; Sequence 59, Application US/10295074
; Publication No. US20030185845A1
; GENERAL INFORMATION:
;  APPLICANT: Pharmexa A/S
;  TITLE OF INVENTION: NOVEL IMMUNOGENIC MIMETICS OF MULTIMER PROTEINS
;  FILE REFERENCE: P1013DK00
;  CURRENT APPLICATION NUMBER: US/10/295,074
;  CURRENT FILING DATE:  2002-11-15
;  NUMBER OF SEQ ID NOS: 60
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 59
;   LENGTH: 514
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: 3 hTNF joined by glycine linkers and P2 and P30 introduced
US-10-295-074-59

  Query Match             58.5%;  Score 117;  DB 14;  Length 514;
```

```
  Best Local Similarity   95.7%;  Pred. No. 1.3e-07;
  Matches   22;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         15 SLFNNFTVSFWLRVPKVSASHLE 37
              :|||||||||||||||||||||||
Db        492 ALFNNFTVSFWLRVPKVSASHLE 514


RESULT 6
US-10-297-942-2
; Sequence 2, Application US/10297942
; Publication No. US20030185816A1
; GENERAL INFORMATION:
;  APPLICANT: Ferring BV
;  TITLE OF INVENTION: Solubilised Protein Vaccines
;  FILE REFERENCE: P68445US0
;  CURRENT APPLICATION NUMBER: US/10/297,942
;  CURRENT FILING DATE:  2003-04-21
;  PRIOR APPLICATION NUMBER: PCT/DK01/00431
;  PRIOR FILING DATE: 2001-10-16
;  PRIOR APPLICATION NUMBER: DK PA 2000 00966
;  PRIOR FILING DATE: 2000-06-21
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 158
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-297-942-2

  Query Match             58.0%;  Score 116;  DB 14;  Length 158;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-08;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         16 LFNNFTVSFWLRVPKVSASHLE 37
              ||||||||||||||||||||||
Db         64 LFNNFTVSFWLRVPKVSASHLE 85


RESULT 7
US-10-295-074-46
; Sequence 46, Application US/10295074
; Publication No. US20030185845A1
; GENERAL INFORMATION:
;  APPLICANT: Pharmexa A/S
;  TITLE OF INVENTION: NOVEL IMMUNOGENIC MIMETICS OF MULTIMER PROTEINS
;  FILE REFERENCE: P1013DK00
;  CURRENT APPLICATION NUMBER: US/10/295,074
;  CURRENT FILING DATE:  2002-11-15
;  NUMBER OF SEQ ID NOS: 60
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 46
;   LENGTH: 194
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: hTNF with inserted tetanus toxoid P2 and P30 epitopes
;   FEATURE:
;   NAME/KEY: MUTAGEN
;   LOCATION: (110)..(124)
;   OTHER INFORMATION: Tetanus toxoid P2 epitope (SEQ ID NO: 2)
;   FEATURE:
;   NAME/KEY: MUTAGEN
;   LOCATION: (125)..(145)
;   OTHER INFORMATION: Tetanus toxoid P30 epitope (SEQ ID NO: 3)
;   FEATURE:
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (2)..(109)
;   OTHER INFORMATION: hTNF amino acids 1-108
;   FEATURE:
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (146)..(194)
;   OTHER INFORMATION: hTNF amino acids 109-157
US-10-295-074-46

  Query Match             58.0%;  Score 116;  DB 14;  Length 194;
  Best Local Similarity   100.0%;  Pred. No. 6.2e-08;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         16 LFNNFTVSFWLRVPKVSASHLE 37
              ||||||||||||||||||||||
Db        124 LFNNFTVSFWLRVPKVSASHLE 145


RESULT 8
US-10-295-074-11
; Sequence 11, Application US/10295074
; Publication No. US20030185845A1
; GENERAL INFORMATION:
;  APPLICANT: Pharmexa A/S
;  TITLE OF INVENTION: NOVEL IMMUNOGENIC MIMETICS OF MULTIMER PROTEINS
;  FILE REFERENCE: P1013DK00
;  CURRENT APPLICATION NUMBER: US/10/295,074
;  CURRENT FILING DATE:  2002-11-15
;  NUMBER OF SEQ ID NOS: 60
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
;   LENGTH: 285
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: 2 human IL5 monomers joined by P2 and P30 epitopes
US-10-295-074-11

  Query Match             58.0%;  Score 116;  DB 14;  Length 285;
  Best Local Similarity   100.0%;  Pred. No. 9.4e-08;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         16 LFNNFTVSFWLRVPKVSASHLE 37
              ||||||||||||||||||||||
Db        149 LFNNFTVSFWLRVPKVSASHLE 170


RESULT 9
US-10-295-074-47
; Sequence 47, Application US/10295074
; Publication No. US20030185845A1
; GENERAL INFORMATION:
;  APPLICANT: Pharmexa A/S
;  TITLE OF INVENTION: NOVEL IMMUNOGENIC MIMETICS OF MULTIMER PROTEINS
;  FILE REFERENCE: P1013DK00
;  CURRENT APPLICATION NUMBER: US/10/295,074
;  CURRENT FILING DATE:  2002-11-15
;  NUMBER OF SEQ ID NOS: 60
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 47
;   LENGTH: 194
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: hTNF with inserted tetanus toxoid P2 and P30 epitopes
;   FEATURE:
;   NAME/KEY: MUTAGEN
;   LOCATION: (110)..(130)
;   OTHER INFORMATION: Tetanus toxoid P30 epitope
;   FEATURE:
;   NAME/KEY: MUTAGEN
;   LOCATION: (131)..(145)
;   OTHER INFORMATION: Tetanus toxoid P2 epitope
;   FEATURE:
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (2)..(109)
;   OTHER INFORMATION: hTNF amino acids 1-108
;   FEATURE:
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (146)..(194)
```

```
;   OTHER INFORMATION: hTNF amino acids 109-157
US-10-295-074-47

  Query Match             57.5%;  Score 115;  DB 14;  Length 194;
  Best Local Similarity   91.7%;  Pred. No. 8.5e-08;
  Matches   22;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         14 PSLFNNFTVSFWLRVPKVSASHLE 37
              |  ||||||||||||||||||||||
Db        107 PEGFNNFTVSFWLRVPKVSASHLE 130


RESULT 10
US-10-452-024-180
; Sequence 180, Application US/10452024
; Publication No. US20040013687A1
; GENERAL INFORMATION:
;  APPLICANT: Simpson, Lance
;  APPLICANT:  Park, Jung-Beak
;  APPLICANT:  Maksymowych, Andrew
;  TITLE OF INVENTION: Compositions and Methods For Transepithelial Molecular Transport
;  FILE REFERENCE: 9855-96U1
;  CURRENT APPLICATION NUMBER: US/10/452,024
;  CURRENT FILING DATE:  2003-06-02
;  PRIOR APPLICATION NUMBER: 60/384,949
;  PRIOR FILING DATE: 2002-05-31
;  NUMBER OF SEQ ID NOS: 188
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 180
;   LENGTH: 441
;   TYPE: PRT
;   ORGANISM: Clostridium botulinum
US-10-452-024-180

  Query Match             57.0%;  Score 114;  DB 15;  Length 441;
  Best Local Similarity   95.5%;  Pred. No. 2.8e-07;
  Matches   21;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         16 LFNNFTVSFWLRVPKVSASHLE 37
              :||||||||||||||||||||||
Db         72 MFNNFTVSFWLRVPKVSASHLE 93


RESULT 11
US-10-452-024-183
; Sequence 183, Application US/10452024
; Publication No. US20040013687A1
; GENERAL INFORMATION:
;  APPLICANT: Simpson, Lance
;  APPLICANT:  Park, Jung-Beak
;  APPLICANT:  Maksymowych, Andrew
;  TITLE OF INVENTION: Compositions and Methods For Transepithelial Molecular Transport
;  FILE REFERENCE: 9855-96U1
;  CURRENT APPLICATION NUMBER: US/10/452,024
;  CURRENT FILING DATE:  2003-06-02
;  PRIOR APPLICATION NUMBER: 60/384,949
;  PRIOR FILING DATE: 2002-05-31
;  NUMBER OF SEQ ID NOS: 188
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 183
;   LENGTH: 441
;   TYPE: PRT
;   ORGANISM: Clostridium botulinum
US-10-452-024-183

  Query Match             57.0%;  Score 114;  DB 15;  Length 441;
  Best Local Similarity   95.5%;  Pred. No. 2.8e-07;
  Matches   21;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         16 LFNNFTVSFWLRVPKVSASHLE 37
              :||||||||||||||||||||||
Db         72 MFNNFTVSFWLRVPKVSASHLE 93


RESULT 12
US-10-452-024-185
; Sequence 185, Application US/10452024
; Publication No. US20040013687A1
; GENERAL INFORMATION:
;  APPLICANT: Simpson, Lance
;  APPLICANT:  Park, Jung-Beak
;  APPLICANT:  Maksymowych, Andrew
;  TITLE OF INVENTION: Compositions and Methods For Transepithelial Molecular Transport
;  FILE REFERENCE: 9855-96U1
;  CURRENT APPLICATION NUMBER: US/10/452,024
;  CURRENT FILING DATE:  2003-06-02
;  PRIOR APPLICATION NUMBER: 60/384,949
;  PRIOR FILING DATE: 2002-05-31
;  NUMBER OF SEQ ID NOS: 188
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 185
;   LENGTH: 444
;   TYPE: PRT
;   ORGANISM: Clostridium tetani
US-10-452-024-185

  Query Match             57.0%;  Score 114;  DB 15;  Length 444;
  Best Local Similarity   95.5%;  Pred. No. 2.8e-07;
  Matches   21;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         16 LFNNFTVSFWLRVPKVSASHLE 37
              :||||||||||||||||||||||
Db         75 MFNNFTVSFWLRVPKVSASHLE 96


RESULT 13
US-10-452-024-186
; Sequence 186, Application US/10452024
; Publication No. US20040013687A1
; GENERAL INFORMATION:
;  APPLICANT: Simpson, Lance
;  APPLICANT:  Park, Jung-Beak
;  APPLICANT:  Maksymowych, Andrew
;  TITLE OF INVENTION: Compositions and Methods For Transepithelial Molecular Transport
;  FILE REFERENCE: 9855-96U1
;  CURRENT APPLICATION NUMBER: US/10/452,024
;  CURRENT FILING DATE:  2003-06-02
;  PRIOR APPLICATION NUMBER: 60/384,949
;  PRIOR FILING DATE: 2002-05-31
;  NUMBER OF SEQ ID NOS: 188
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 186
;   LENGTH: 451
;   TYPE: PRT
;   ORGANISM: Clostridium tetani
US-10-452-024-186

  Query Match             57.0%;  Score 114;  DB 15;  Length 451;
  Best Local Similarity   95.5%;  Pred. No. 2.9e-07;
  Matches   21;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         16 LFNNFTVSFWLRVPKVSASHLE 37
              :||||||||||||||||||||||
Db         82 MFNNFTVSFWLRVPKVSASHLE 103


RESULT 14
US-10-452-024-184
; Sequence 184, Application US/10452024
; Publication No. US20040013687A1
; GENERAL INFORMATION:
;  APPLICANT: Simpson, Lance
;  APPLICANT:  Park, Jung-Beak
;  APPLICANT:  Maksymowych, Andrew
;  TITLE OF INVENTION: Compositions and Methods For Transepithelial Molecular Transport
;  FILE REFERENCE: 9855-96U1
```

; CURRENT APPLICATION NUMBER: US/10/452,024
; CURRENT FILING DATE: 2003-06-02
; PRIOR APPLICATION NUMBER: 60/384,949
; PRIOR FILING DATE: 2002-05-31
; NUMBER OF SEQ ID NOS: 188
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 184
; LENGTH: 452
; TYPE: PRT
; ORGANISM: Clostridium botulinum
US-10-452-024-184

Query Match 57.0%; Score 114; DB 15; Length 452;
Best Local Similarity 95.5%; Pred. No. 2.9e-07;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LFNNFTVSFWLRVPKVSASHLE 37
:||||||||||||||||||||||
Db 83 MFNNFTVSFWLRVPKVSASHLE 104

RESULT 15
US-09-816-467-2
; Sequence 2, Application US/09816467
; Publication No. US20030004121A1
; GENERAL INFORMATION:
; APPLICANT: COEN, LAURENT
; APPLICANT: PINZOLAS, ROSARIO OSTA
; APPLICANT: BRULET, PHILIPPE
; TITLE OF INVENTION: HYBRID PROTEINS THAT MIGRATE RETROGRADELY AND
; TITLE OF INVENTION: TRANSYNAPTICALLY INTO THE CNS
; FILE REFERENCE: 03495.0174-01000
; CURRENT APPLICATION NUMBER: US/09/816,467
; CURRENT FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: 60/055,615
; PRIOR FILING DATE: 1997-08-14
; PRIOR APPLICATION NUMBER: 60/065,236
; PRIOR FILING DATE: 1997-11-13
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 463
; TYPE: PRT
; ORGANISM: Clostridium tetani
US-09-816-467-2

Query Match 57.0%; Score 114; DB 10; Length 463;
Best Local Similarity 95.5%; Pred. No. 3e-07;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LFNNFTVSFWLRVPKVSASHLE 37
:||||||||||||||||||||||
Db 94 MFNNFTVSFWLRVPKVSASHLE 115

Search completed: March 10, 2004, 10:25:49
Job time : 30.0817 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:48 ; Search time 55.716 Seconds
(without alignments)
187.635 Million cell updates/sec

Title: US-09-848-834A-14
Perfect score: 200
Sequence: 1 XHWSYGLRPGSSGPSLFNNFTVSFWLRVPKVSASHLE 37

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	199	99.5	37	5	AAU11425	Aau11425 Synthetic
2	199	99.5	50	5	AAU11429	Aau11429 Synthetic
3	123	61.5	109	4	AAB20150	Aab20150 Growth di
4	121	60.5	216	3	AAY92665	Aay92665 MUC-1 ana
5	119.5	59.8	109	4	AAB20151	Aab20151 Growth di
6	119.5	59.8	750	3	AAY92644	Aay92644 Mutant hu
7	119	59.5	122	3	AAB45524	Aab45524 Modified
8	119	59.5	122	3	AAB45507	Aab45507 Modified
9	118.5	59.2	158	2	AAW81332	Aaw81332 TNF30-1,
10	118.5	59.2	158	5	ABB07282	Abb07282 Human TNF
11	118	59.0	109	4	AAB20149	Aab20149 Growth di
12	118	59.0	703	3	AAY92662	Aay92662 Mutant mu
13	118	59.0	761	3	AAY92660	Aay92660 Mutant mu
14	117	58.5	124	3	AAB45496	Aab45496 Modified
15	117	58.5	124	3	AAB45515	Aab45515 Modified
16	117	58.5	160	4	AAB20153	Aab20153 Growth di
17	117	58.5	287	6	AAO30459	Aao30459 hIL5.36 v
18	117	58.5	514	6	AAO30495	Aao30495 Human TNF
19	117	58.5	708	7	ABR82479	Abr82479 Modified
20	116	58.0	31	3	AAY92653	Aay92653 PSMpep010
21	116	58.0	43	4	AAB49076	Aab49076 Amyloid b
22	116	58.0	43	4	AAB46177	Aab46177 Tetanus t
23	116	58.0	72	4	AAB46190	Aab46190 Tetanus t
24	116	58.0	109	4	AAB20148	Aab20148 Growth di
25	116	58.0	136	4	AAB49089	Aab49089 Amyloid b
26	116	58.0	145	3	AAB45530	Aab45530 Modified
27	116	58.0	147	3	AAB45522	Aab45522 Modified
28	116	58.0	158	2	AAW81334	Aaw81334 TNF30-3,
29	116	58.0	158	5	ABB07274	Abb07274 Human TNF
30	116	58.0	188	3	AAY84423	Aay84423 An osteop
31	116	58.0	194	6	AAO30488	Aao30488 Human TNF
32	116	58.0	254	4	AAB20152	Aab20152 Growth di
33	116	58.0	285	6	AAO30458	Aao30458 hIL5-P2-P
34	116	58.0	717	7	ABR82478	Abr82478 Modified
35	116	58.0	750	3	AAY92637	Aay92637 Mutant hu
36	116	58.0	750	3	AAY92639	Aay92639 Mutant hu
37	116	58.0	750	3	AAY92628	Aay92628 Mutant hu
38	116	58.0	750	3	AAY92631	Aay92631 Mutant hu
39	116	58.0	750	3	AAY92627	Aay92627 Mutant hu
40	116	58.0	750	3	AAY92638	Aay92638 Mutant hu
41	116	58.0	750	3	AAY92630	Aay92630 Mutant hu
42	116	58.0	750	3	AAY92629	Aay92629 Mutant hu
43	116	58.0	750	3	AAY92642	Aay92642 Mutant hu
44	115	57.5	194	6	AAO30489	Aao30489 Human TNF
45	115	57.5	693	3	AAY92647	Aay92647 Mutant hu

ALIGNMENTS

```
RESULT 1
AAU11425
ID   AAU11425 standard; peptide; 37 AA.
XX
AC   AAU11425;
XX
DT   12-MAR-2002  (first entry)
XX
DE   Synthetic immunogen peptide 6.
XX
KW   Gonadotrophin releasing hormone; GnRH; synthetic immunogen;
KW   luteinising hormone releasing hormone; LHRH; contraceptive;
KW   promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
KW   breast cancer; uterine cancer; gynaecological cancer; endometriosis;
KW   uterine fibroid; benign prostatic hypertrophy; prostate cancer.
XX
OS   Clostridium tetani.
OS   Mammalia.
OS   Synthetic.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .10
FT                   /note= "Gonadotrophin releasing hormone epitope"
FT   Misc-difference 1
FT                   /label= OTHER
FT                   /note= "Other= Pyro-glutamic acid or 5-oxo proline"
FT   Peptide         11. .16
FT                   /note= "Spacer peptide"
FT   Peptide         17. .37
FT                   /note= "Tetanus toxoid sequence (947-967 aa)"
XX
PN   WO200185763-A2.
XX
PD   15-NOV-2001.
XX
PF   04-MAY-2001; 2001WO-US014363.
XX
PR   05-MAY-2000; 2000US-0202328P.
XX
PA   (APHT-) APHTON CORP.
XX
PI   Grimes S,  Michaeli D,  Stevens VC;
XX
DR   WPI; 2002-049440/06.
XX
PT   Novel synthetic immunogen for inducing immune response against
PT   gonadotropin releasing hormone, comprises fusion peptide having
```

```
PT   promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
PT   or its analog.
XX
PS   Claim 11; Page 9; 43pp; English.
XX
CC   The invention relates to a synthetic immunogen for inducing specific
CC   antibodies against gonadotropin releasing hormone (GnRH also known as
CC   luteinising hormone releasing hormone, LHRH) comprising a fusion peptide
CC   which comprises a promiscuous helper T-cell peptide epitope and
CC   immunomimic peptide epitope or its analogue. The synthetic immunogen is
CC   useful inducing an immune response against GnRH in an animal subject, and
CC   as such is useful as a contraceptive and in the treatment of diseases
CC   such as cancer (of the breast, uterus and other gynaecological cancer),
CC   endometriosis, uterine fibroids, benign prostatic hypertrophy and
CC   prostate cancer. The immunogen is effective in eliciting high and
CC   specific anti-GnRH antibody titres. The present sequence is a synthetic
CC   immunogen of the invention
XX
SQ   Sequence 37 AA;

  Query Match             99.5%;  Score 199;  DB 5;  Length 37;
  Best Local Similarity   100.0%;  Pred. No. 1e-21;
  Matches   36;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 HWSYGLRPGSSGPSLFNNFTVSFWLRVPKVSASHLE 37
              |||||||||||||||||||||||||||||||||||
Db          2 HWSYGLRPGSSGPSLFNNFTVSFWLRVPKVSASHLE 37


RESULT 2
AAU11429
ID   AAU11429 standard; peptide; 50 AA.
XX
AC   AAU11429;
XX
DT   12-MAR-2002  (first entry)
XX
DE   Synthetic immunogen peptide 10.
XX
KW   Gonadotrophin releasing hormone; GnRH; synthetic immunogen;
KW   luteinising hormone releasing hormone; LHRH; contraceptive;
KW   promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
KW   breast cancer; uterine cancer; gynaecological cancer; endometriosis;
KW   uterine fibroid; benign prostatic hypertrophy; prostate cancer.
XX
OS   Clostridium tetani.
OS   Mammalia.
OS   Synthetic.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .10
FT                   /note= "Gonadotrophin releasing hormone epitope (1. .10
FT                   aa)"
FT   Misc-difference 1
FT                   /label= OTHER
FT                   /note= "Other= Pyro-glutamic acid or 5-oxo proline"
FT   Peptide         11. .16
FT                   /note= "Spacer peptide"
FT   Peptide         17. .37
FT                   /note= "Tetanus toxoid (947-967 aa)"
FT   Peptide         38. .41
FT                   /note= "Spacer peptide"
FT   Peptide         42. .50
FT                   /note= "Gonadotrophin releasing hormone epitope (2-10
FT                   aa)"
FT   Modified-site   50
FT                   /note= "Amidated glycine or glycinamide"
XX
PN   WO200185763-A2.
XX
PD   15-NOV-2001.
XX
PF   04-MAY-2001; 2001WO-US014363.
XX
PR   05-MAY-2000; 2000US-0202328P.
XX
PA   (APHT-) APHTON CORP.
XX
PI   Grimes S,  Michaeli D,  Stevens VC;
XX
DR   WPI; 2002-049440/06.
XX
PT   Novel synthetic immunogen for inducing immune response against
PT   gonadotropin releasing hormone, comprises fusion peptide having
PT   promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
PT   or its analog.
XX
PS   Claim 11; Page 11; 43pp; English.
XX
CC   The invention relates to a synthetic immunogen for inducing specific
CC   antibodies against gonadotropin releasing hormone (GnRH also known as
CC   luteinising hormone releasing hormone, LHRH) comprising a fusion peptide
CC   which comprises a promiscuous helper T-cell peptide epitope and
CC   immunomimic peptide epitope or its analogue. The synthetic immunogen is
CC   useful inducing an immune response against GnRH in an animal subject, and
CC   as such is useful as a contraceptive and in the treatment of diseases
CC   such as cancer (of the breast, uterus and other gynaecological cancer),
CC   endometriosis, uterine fibroids, benign prostatic hypertrophy and
CC   prostate cancer. The immunogen is effective in eliciting high and
CC   specific anti-GnRH antibody titres. The present sequence is a synthetic
CC   immunogen of the invention
XX
SQ   Sequence 50 AA;

  Query Match             99.5%;  Score 199;  DB 5;  Length 50;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-21;
  Matches   36;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 HWSYGLRPGSSGPSLFNNFTVSFWLRVPKVSASHLE 37
              |||||||||||||||||||||||||||||||||||
Db          2 HWSYGLRPGSSGPSLFNNFTVSFWLRVPKVSASHLE 37


RESULT 3
AAB20150
ID   AAB20150 standard; protein; 109 AA.
XX
AC   AAB20150;
XX
DT   30-APR-2001  (first entry)
XX
DE   Growth differentiation factor 8 AutoVac construct GDF-8 P30-3A.
XX
KW   Growth differentiation factor 8; GDF-8; myostatin; tetanus toxin;
KW   T-cell epitope; down-regulation; vaccine; muscle; meat; cachexia;
KW   cardiant; human; mutant; mutein.
XX
OS   Homo sapiens.
OS   Clostridium tetani.
OS   Synthetic.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   Region          1. .78
FT                   /note= "identical to residues 267-345 of human GDF-8"
FT   Misc-difference 73
FT                   /note= "Cys-73 may be substituted by Ser to avoid
FT                   disulfide bond formation"
FT   Region          79. .99
FT                   /note= "tetanus toxoid P2 epitope"
FT   Misc-difference 90. .91
FT                   /note= "optionally replaced by Glu-Gly"
FT   Region          100. .109
FT                   /note= "identical to residues 366-375 of human GDF-8"
XX
```

PN WO200105820-A2.
XX
PD 25-JAN-2001.
XX
PF 20-JUL-2000; 2000WO-DK000413.
XX
PR 20-JUL-1999; 99DK-00001014.
PR 26-JUL-1999; 99US-0145275P.
XX
PA (MEBI-) M & E BIOTECH AS.
XX
PI Halkier T, Mouritsen S, Klysner S;
XX
DR WPI; 2001-112680/12.
XX
PT Increasing the muscle mass of animals used in meat production by down
PT regulating growth differentiation factor 8 (GDF-8) activity in the animal
PT through induction of anti-GDF-8 antibody production.
XX
PS Example 1; Page 102-103; 110pp; English.
XX
CC The present sequence is that of AutoVac construct GDF-8 P30-3A,
CC comprising the 109 C-terminal amino acid residues of human growth
CC differentiation factor 8 (GDF-8) in which residues 79-99 are replaced by
CC the promiscuous tetanus toxin T-cell epitope P30 (see AAB20144). It is an
CC object of the invention to produce a recombinant therapeutic vaccine that
CC is capable of effecting down-regulation of GDF-8 in order to increase the
CC muscle growth rate of farm animals. The vaccines (see AAB20145-53) are
CC capable of breaking autotolerance against autologous GDF-8. They comprise
CC the C-terminal portion of human GDF-8 in which a portion of the native
CC sequence is replaced by a T-cell epitope such as P30, with minimal
CC disturbance of the authentic 3-dimensional structure of the protein.
CC Nucleic acids encoding the GDF-8 variants can be used for genetic
CC immunisation of the animals. Down-regulation of GDF-8 activity can
CC increase muscle mass by up to at least 45% in cattle, pigs and poultry
CC used for meat production, reducing the need for antibiotic feed-
CC additives. Anti-GDF8 vaccines can be used to treat human diseases such as
CC cancer cachexia where muscle atrophy is pronounced and for patients
CC suffering from acute and chronic heart failure
XX
SQ Sequence 109 AA;

Query Match 61.5%; Score 123; DB 4; Length 109;
Best Local Similarity 78.1%; Pred. No. 5.5e-10;
Matches 25; Conservative 2; Mismatches 1; Indels 4; Gaps 1;

```
Qy         10 GSSG----PSLFNNFTVSFWLRVPKVSASHLE 37
              ||:|    |: ||||||||||||||||||||||
Db         68 GSAGPCCTPTKFNNFTVSFWLRVPKVSASHLE 99
```

RESULT 4
AAY92665
ID AAY92665 standard; peptide; 216 AA.
XX
AC AAY92665;
XX
DT 10-AUG-2000 (first entry)
XX
DE MUC-1 analogue containing foreign epitopes.
XX
KW Mucin repeat; MUC-1 analogue; vaccination; self-protein; cancer;
KW cytotoxic T-lymphocyte immunity; breast cancer; prostate cancer;
KW cell-associated peptide antigen; foreign epitope.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 61. .75
FT /label= P2
FT Peptide 136. .156
FT /label= P30
FT /note= "q"
XX
PN WO200020027-A2.
XX
PD 13-APR-2000.
XX
PF 05-OCT-1999; 99WO-DK000525.
XX
PR 05-OCT-1998; 98DK-00001261.
PR 20-OCT-1998; 98US-0105011P.
XX
PA (MEBI-) M & E BIOTECH AS.
XX
PI Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
PI Gautam A, Birk P, Karlsson G;
XX
DR WPI; 2000-349917/30.
XX
PT Inducing immune responses to weakly immunogenic, tumor associated peptide
PT antigens for the treatment of breast and prostate cancer.
XX
PS Example 4; Page; 220pp; English.
XX
CC This is an immunogenized MUC-1 analogue containing foreign epitopes P2
CC and P30. Immunogenic analogues of MUC-1 and, e.g. human prostate specific
CC membrane antigen (hPSM) can be used in the claimed method as an
CC autovaccine to induce a CTL response. Subdominant CTL epitopes, antibody
CC binding regions and cysteine residues involved in disulfide bonds are
CC preserved in the immunogenized forms (see features table). 10 regions
CC suitable for the insertion of foreign T helper epitopes were identified.
CC The method is used for inducing immune responses against weakly
CC immunogenic cell-associated peptide antigens (PA) such as those
CC associated with cancers (self-proteins), e.g. hPSM, heregulin 2 (Her2)
CC and/or fibroblast growth factor 8b (FGF8b). The method comprises
CC effecting simultaneous presentation by antigen producing cells (APCs) of
CC the animals immune system of: (1) at least 1 CTL (cytotoxic T-lymphocyte)
CC group derived from the PA and/or at least 1 B-cell group derived from the
CC cell-associated PA; and (2) at least 1 first T helper cell group which is
CC foreign to the animal. Analogues of human PSM, human Her2 and
CC human/murine FGF8b comprising a substantial part of all known and
CC predicted CTL and B-cell epitopes of the respective PA and including at
CC least one foreign T helper epitope are also claimed. The method is used
CC to treat prostate, prostate/breast or breast cancer when the PA is human
CC PSM, FGF8b and Her2, respectively. Note: This sequence does not appear in
CC the specification. It was made using the mucin repeat sequence
CC (AAY92664), P2 and P30 (AAY92625-26), which appear on pages 220, 213 and
CC 214 respectively, of the specification
XX
SQ Sequence 216 AA;

Query Match 60.5%; Score 121; DB 3; Length 216;
Best Local Similarity 64.1%; Pred. No. 2.4e-09;
Matches 25; Conservative 2; Mismatches 2; Indels 10; Gaps 1;

```
Qy          9 PGSSGP----------SLFNNFTVSFWLRVPKVSASHLE 37
              |||: |          : ||||||||||||||||||||||
Db        118 PGSTAPPAHGVTSAPDTRFNNFTVSFWLRVPKVSASHLE 156
```

RESULT 5
AAB20151
ID AAB20151 standard; protein; 109 AA.
XX
AC AAB20151;
XX
DT 30-APR-2001 (first entry)
XX
DE Growth differentiation factor 8 AutoVac construct GDF-8 P30-3B.
XX
KW Growth differentiation factor 8; GDF-8; myostatin; tetanus toxin;
KW T-cell epitope; down-regulation; vaccine; muscle; meat; cachexia;
KW cardiant; human; mutant; mutein.
XX
OS Homo sapiens.

OS Clostridium tetani.
OS Synthetic.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT Region 1. .83
FT /note= "identical to residues 267-349 of human GDF-8"
FT Misc-difference 73
FT /note= "Cys-73 may be substituted by Ser to avoid
FT disulfide bond formation"
FT Region 84. .104
FT /note= "tetanus toxoid P2 epitope"
FT Misc-difference 90. .91
FT /note= "optionally replaced by Glu-Gly"
FT Region 105. .109
FT /note= "identical to residues 371-375 of human GDF-8"
XX
PN WO200105820-A2.
XX
PD 25-JAN-2001.
XX
PF 20-JUL-2000; 2000WO-DK000413.
XX
PR 20-JUL-1999; 99DK-00001014.
PR 26-JUL-1999; 99US-0145275P.
XX
PA (MEBI-) M & E BIOTECH AS.
XX
PI Halkier T, Mouritsen S, Klysner S;
XX
DR WPI; 2001-112680/12.
XX
PT Increasing the muscle mass of animals used in meat production by down
PT regulating growth differentiation factor 8 (GDF-8) activity in the animal
PT through induction of anti-GDF-8 antibody production.
XX
PS Example 1; Page 104; 110pp; English.
XX
CC The present sequence is that of AutoVac construct GDF-8 P30-3B,
CC comprising the 109 C-terminal amino acid residues of human growth
CC differentiation factor 8 (GDF-8) in which residues 84-104 are replaced by
CC the promiscuous tetanus toxin T-cell epitope P30 (see AAB20144). It is an
CC object of the invention to produce a recombinant therapeutic vaccine that
CC is capable of effecting down-regulation of GDF-8 in order to increase the
CC muscle growth rate of farm animals. The vaccines (see AAB20145-53) are
CC capable of breaking autotolerance against autologous GDF-8. They comprise
CC the C-terminal portion of human GDF-8 in which a portion of the native
CC sequence is replaced by a T-cell epitope such as P30, with minimal
CC disturbance of the authentic 3-dimensional structure of the protein.
CC Nucleic acids encoding the GDF-8 variants can be used for genetic
CC immunisation of the animals. Down-regulation of GDF-8 activity can
CC increase muscle mass by up to at least 45% in cattle, pigs and poultry
CC used for meat production, reducing the need for antibiotic feed-
CC additives. Anti-GDF8 vaccines can be used to treat human diseases such as
CC cancer cachexia where muscle atrophy is pronounced and for patients
CC suffering from acute and chronic heart failure
XX
SQ Sequence 109 AA;

Query Match 59.8%; Score 119.5; DB 4; Length 109;
Best Local Similarity 67.6%; Pred. No. 1.8e-09;
Matches 25; Conservative 1; Mismatches 2; Indels 9; Gaps 1;

Qy 10 GSSGPSL---------FNNFTVSFWLRVPKVSASHLE 37
||:|| |||||||||||||||||||||
Db 68 GSAGPCCTPTKMSPINFNNFTVSFWLRVPKVSASHLE 104

RESULT 6
AAY92644
ID AAY92644 standard; protein; 750 AA.
XX
AC AAY92644;
XX
DT 10-AUG-2000 (first entry)
XX
DE Mutant human prostate specific membrane antigen construct, hPSM6.3.
XX
KW Prostate specific membrane antigen; immunogenized construct; mutant;
KW vaccination; cytotoxic T-lymphocyte immunity; breast cancer;
KW prostate cancer; cell-associated peptide antigen; foreign epitope.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Peptide 210. .230
FT /label= P30
FT /note= "foreign epitope"
FT Peptide 448. .462
FT /label= P2
FT /note= "foreign epitope"
XX
PN WO200020027-A2.
XX
PD 13-APR-2000.
XX
PF 05-OCT-1999; 99WO-DK000525.
XX
PR 05-OCT-1998; 98DK-00001261.
PR 20-OCT-1998; 98US-0105011P.
XX
PA (MEBI-) M & E BIOTECH AS.
XX
PI Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
PI Gautam A, Birk P, Karlsson G;
XX
DR WPI; 2000-349917/30.
XX
PT Inducing immune responses to weakly immunogenic, tumor associated peptide
PT antigens for the treatment of breast and prostate cancer.
XX
PS Example 1; Page; 220pp; English.
XX
CC AAY92627-49 are mutant immunogenized human prostate specific membrane
CC antigen (PSM) constructs, which contain foreign epitopes (P2 and/or P30).
CC The immunogenic analogues of PSM can be used in the claimed method as an
CC autovaccine to induce a CTL response. Subdominant CTL epitopes, antibody
CC binding regions and cysteine residues involved in disulfide bonds are
CC preserved in the immunogenized forms. The method is used for inducing
CC immune responses against weakly immunogenic cell-associated peptide
CC antigens (PA) such as those associated with cancers (self-proteins), e.g.
CC human prostate specific membrane antigen (PSM), heregulin 2 (Her2) and/or
CC fibroblast growth factor 8b (FGF8b). The method comprises effecting
CC simultaneous presentation by antigen producing cells (APCs) of the
CC animals immune system of: (1) at least 1 CTL (cytotoxic T-lymphocyte)
CC group derived from the PA and/or at least 1 B-cell group derived from the
CC cell-associated PA; and (2) at least 1 first T helper cell group which is
CC foreign to the animal. Analogues of human PSM, human Her2 and
CC human/murine FGF8b comprising a substantial part of all known and
CC predicted CTL and B-cell epitopes of the respective PA and including at
CC least one foreign T helper epitope are also claimed. The method is used
CC to treat prostate, prostate/breast or breast cancer when the PA is human
CC PSM, FGF8b and Her2, respectively. Note: This sequence was constructed
CC from the wild type human PSM (AAY92619), which appears on pages 184-187
CC of the specification
XX
SQ Sequence 750 AA;

Query Match 59.8%; Score 119.5; DB 3; Length 750;
Best Local Similarity 78.1%; Pred. No. 1.7e-08;
Matches 25; Conservative 0; Mismatches 4; Indels 3; Gaps 1;

Qy 9 PGSS---GPSLFNNFTVSFWLRVPKVSASHLE 37
|| | ||||||||||||||||||||||
Db 249 PGGGVQRGNILFNNFTVSFWLRVPKVSASHLE 280

```
RESULT 7
AAB45524
ID   AAB45524 standard; protein; 122 AA.
XX
AC   AAB45524;
XX
DT   26-FEB-2001  (first entry)
XX
DE   Modified murine interleukin-5 SEQ ID NO: 48.
XX
KW   Asthma; IL-5; interleukin-5; allergy; cytokine; helminthic infection;
KW   cancer; eosinophilia; vaccine; allergic rhinitis.
XX
OS   Mus musculus.
OS   Clostridium tetani.
XX
PN   WO200065058-A1.
XX
PD   02-NOV-2000.
XX
PF   19-APR-2000; 2000WO-DK000205.
XX
PR   23-APR-1999;   99DK-00000552.
PR   06-MAY-1999;   99US-0132811P.
XX
PA   (MEBI-) M & E BIOTECH AS.
XX
PI   Klysner S;
XX
DR   WPI; 2000-672791/65.
DR   N-PSDB; AAC68877.
XX
PT   Down-regulating interleukin 5 (IL-5) activity in humans by administering
PT   IL-5 and/or an IL-5 analogue, useful in the treatment, prophylaxis or
PT   amelioration of asthma or other chronic allergic conditions.
XX
PS   Example 7; Page 156; 172pp; English.
XX
CC   The present invention is concerned with methods of treating asthma,
CC   eosinophilia, allergic rhinitis and other allergic diseases. These
CC   involve the use of interleukin-5 (IL-5) analogues and modified IL-5
CC   proteins and their coding sequences to down-regulate IL-5 activity and
CC   thus reduce eosinophil numbers. The allergic diseases may be treated
CC   using autovaccines, nucleic acid vaccines or live vaccines. In addition,
CC   it is possible that they may be used in the treatment of cancer and
CC   helminthic infections
XX
SQ   Sequence 122 AA;

  Query Match             59.5%;  Score 119;  DB 3;  Length 122;
  Best Local Similarity   81.5%;  Pred. No. 2.4e-09;
  Matches   22;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         11 SSGPSLFNNFTVSFWLRVPKVSASHLE 37
              :|  ::|||||||||||||||||||||
Db         24 TSNETMFNNFTVSFWLRVPKVSASHLE 50


RESULT 8
AAB45507
ID   AAB45507 standard; protein; 122 AA.
XX
AC   AAB45507;
XX
DT   26-FEB-2001  (first entry)
XX
DE   Modified murine interleukin-5 SEQ ID NO: 19.
XX
KW   Asthma; IL-5; interleukin-5; allergy; cytokine; helminthic infection;
KW   cancer; eosinophilia; vaccine; allergic rhinitis.
XX
OS   Mus musculus.
OS   Clostridium tetani.
XX
PN   WO200065058-A1.
XX
PD   02-NOV-2000.
XX
PF   19-APR-2000; 2000WO-DK000205.
XX
PR   23-APR-1999;   99DK-00000552.
PR   06-MAY-1999;   99US-0132811P.
XX
PA   (MEBI-) M & E BIOTECH AS.
XX
PI   Klysner S;
XX
DR   WPI; 2000-672791/65.
XX
PT   Down-regulating interleukin 5 (IL-5) activity in humans by administering
PT   IL-5 and/or an IL-5 analogue, useful in the treatment, prophylaxis or
PT   amelioration of asthma or other chronic allergic conditions.
XX
PS   Example 7; Page 134; 172pp; English.
XX
CC   The present invention is concerned with methods of treating asthma,
CC   eosinophilia, allergic rhinitis and other allergic diseases. These
CC   involve the use of interleukin-5 (IL-5) analogues and modified IL-5
CC   proteins and their coding sequences to down-regulate IL-5 activity and
CC   thus reduce eosinophil numbers. The allergic diseases may be treated
CC   using autovaccines, nucleic acid vaccines or live vaccines. In addition,
CC   it is possible that they may be used in the treatment of cancer and
CC   helminthic infections
XX
SQ   Sequence 122 AA;

  Query Match             59.5%;  Score 119;  DB 3;  Length 122;
  Best Local Similarity   81.5%;  Pred. No. 2.4e-09;
  Matches   22;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         11 SSGPSLFNNFTVSFWLRVPKVSASHLE 37
              :|  ::|||||||||||||||||||||
Db         24 TSNETMFNNFTVSFWLRVPKVSASHLE 50


RESULT 9
AAW81332
ID   AAW81332 standard; protein; 158 AA.
XX
AC   AAW81332;
XX
DT   21-APR-1999  (first entry)
XX
DE   TNF30-1, a TNF-alpha analogue.
XX
KW   Human tumour necrosis factor-alpha; TNF-alpha; TNF-alpha analogue;
KW   vaccine; rheumatoid arthritis; Crohn's disease; ulcerative colitis;
KW   cancer; disseminated sclerosis; diabetes; psoriasis; osteoporosis;
KW   asthma.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9846642-A1.
XX
PD   22-OCT-1998.
XX
PF   15-APR-1998;   98WO-DK000157.
XX
PR   15-APR-1997;   97DK-00000418.
PR   24-APR-1997;   97US-0044187P.
XX
PA   (FERR ) FARM LAB FERRING AS.
XX
```

PI Jensen MR, Mouritsen S, Elsner H, Dalum I;
XX
DR WPI; 1998-594561/50.
DR N-PSDB; AAV68421.
XX
PT Modified human tumour necrosis factor-alpha - comprises immunodominant T
PT cell epitope, useful in vaccines to treat or prevent TNF-associated
PT diseases, e.g. cancer.
XX
PS Example 1; Page 74-75; 134pp; English.
XX
CC The present sequence represents a modified human tumour necrosis factor-
CC alpha (TNF-alpha) analogue. The analogues have no residual TNF activity
CC and are immunogenic in a large proportion of the human population (by
CC using promiscuous epitopes). The TNF-alpha analogue is able to generate,
CC in humans, neutralizing antibodies to wild-type human TNF alpha, has at
CC least one fragment of TNF substituted by a peptide containing an
CC immunodominant T-cell epitope, and at least one TNF-alpha B-cell epitope.
CC The substitution causes a significant change in the amino acid sequence
CC of any one of the strands in the front beta-sheet, any of the connecting
CC loops or any of the B', I or D strands in the back beta-sheet. The TNF-
CC alpha analogues are used as vaccines for treatment or prevention of
CC diseases associated with excessive release or activity of TNF-alpha, e.g.
CC rheumatoid arthritis, Crohn's disease, ulcerative colitis, cancer of any
CC sort, disseminated sclerosis, diabetes, psoriasis, osteoporosis and
CC asthma
XX
SQ Sequence 158 AA;

Query Match 59.2%; Score 118.5; DB 2; Length 158;
Best Local Similarity 80.6%; Pred. No. 3.9e-09;
Matches 25; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

Qy 7 LRPGSSGPSLFNNFTVSFWLRVPKVSASHLE 37
:| | || ||||||||||||||||||||||
Db 2 VRSSSRTPS-FNNFTVSFWLRVPKVSASHLE 31

RESULT 10
ABB07282
ID ABB07282 standard; protein; 158 AA.
XX
AC ABB07282;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human TNF-alpha analogue TNF30-1.
XX
KW TNF-alpha; pharmaceutical; vaccine; self-protein; tumour necrosis factor;
KW cetylpyridinium chloride; immunisation; antiinflammatory; antirheumatic;
KW antiarthritic; antiulcer; cytostatic; antidiabetic; antipsoriatic;
KW antiasthmatic; immunomodulator; neuroprotective; osteopathic; human;
KW TNF30-1.
XX
OS Homo sapiens.
XX
PN WO200197837-A1.
XX
PD 27-DEC-2001.
XX
PF 20-JUN-2001; 2001WO-DK000431.
XX
PR 21-JUN-2000; 2000DK-00000966.
XX
PA (FERR) FERRING BV.
XX
PI Olesen OF, Balchen T, Bouman MHEM;
XX
DR WPI; 2002-114542/15.
DR N-PSDB; ABA94392.
XX
PT Novel vaccine composition for prevention/treatment of self-protein-
PT mediated pathology such as cancer, diabetes and asthma, comprises
PT modified immunogenic self-protein and surfactant capable of acting as
PT solubilizer.
XX
PS Claim 21; Page 48-49; 55pp; English.
XX
CC The invention provides a pharmaceutical vaccine composition (I) for the
CC prevention or treatment of a self-protein-mediated pathology. The
CC composition comprises at least one modified immunogenic self-protein
CC (selected from modified TNF-alpha proteins) and a surfactant capable of
CC acting as a solubilizer. (I) is useful for preventing or treating a self
CC -protein-mediated pathology such as an inflammatory disease, rheumatoid
CC arthritis, an inflammatory bowel disease (ulcerative colitis or Crohn's
CC disease), cancer, cachexia, multiple sclerosis, diabetes, psoriasis,
CC osteoporosis or asthma. (I) is useful for inducing autoantibodies to a
CC self-protein such as TNF (tumour necrosis factor)-alpha in a human
CC subject. (I) comprising cetylpyridinium chloride as a component is useful
CC for immunisation of a human subject or for treatment of a human
CC inflammatory disease. The present sequence represents a human TNF-alpha
CC analogue TNF30-1
XX
SQ Sequence 158 AA;

Query Match 59.2%; Score 118.5; DB 5; Length 158;
Best Local Similarity 80.6%; Pred. No. 3.9e-09;
Matches 25; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

Qy 7 LRPGSSGPSLFNNFTVSFWLRVPKVSASHLE 37
:| | || ||||||||||||||||||||||
Db 2 VRSSSRTPS-FNNFTVSFWLRVPKVSASHLE 31

RESULT 11
AAB20149
ID AAB20149 standard; protein; 109 AA.
XX
AC AAB20149;
XX
DT 30-APR-2001 (first entry)
XX
DE Growth differentiation factor 8 AutoVac construct GDF-8 P30-2.
XX
KW Growth differentiation factor 8; GDF-8; myostatin; tetanus toxin;
KW T-cell epitope; down-regulation; vaccine; muscle; meat; cachexia;
KW cardiant; human; mutant; mutein.
XX
OS Homo sapiens.
OS Clostridium tetani.
OS Synthetic.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT Region 1. .48
FT /note= "identical to residues 267-314 of human GDF-8"
FT Region 49. .69
FT /note= "tetanus toxoid P2 epitope"
FT Region 70. .109
FT /note= "identical to residues 336-375 of human GDF-8"
FT Misc-difference 73
FT /note= "Cys-73 may be substituted by Ser to avoid
FT disulfide bond formation"
FT Misc-difference 90. .91
FT /note= "optionally replaced by Glu-Gly"
XX
PN WO200105820-A2.
XX
PD 25-JAN-2001.
XX
PF 20-JUL-2000; 2000WO-DK000413.
XX
PR 20-JUL-1999; 99DK-00001014.
PR 26-JUL-1999; 99US-0145275P.
XX
PA (MEBI-) M & E BIOTECH AS.

XX
PI Halkier T, Mouritsen S, Klysner S;
XX
DR WPI; 2001-112680/12.
XX
PT Increasing the muscle mass of animals used in meat production by down
PT regulating growth differentiation factor 8 (GDF-8) activity in the animal
PT through induction of anti-GDF-8 antibody production.
XX
PS Example 1; Page 101-102; 110pp; English.
XX
CC The present sequence is that of AutoVac construct GDF-8 P30-2, comprising
CC the 109 C-terminal amino acid residues of human growth differentiation
CC factor 8 (GDF-8) in which residues 49-69 are replaced by the promiscuous
CC tetanus toxin T-cell epitope P30 (see AAB20144). It is an object of the
CC invention to produce a recombinant therapeutic vaccine that is capable of
CC effecting down-regulation of GDF-8 in order to increase the muscle growth
CC rate of farm animals. The vaccines (see AAB20145-53) are capable of
CC breaking autotolerance against autologous GDF-8. They comprise the C-
CC terminal portion of human GDF-8 in which a portion of the native sequence
CC is replaced by a T-cell epitope such as P30, with minimal disturbance of
CC the authentic 3-dimensional structure of the protein. Nucleic acids
CC encoding the GDF-8 variants can be used for genetic immunisation of the
CC animals. Down-regulation of GDF-8 activity can increase muscle mass by up
CC to at least 45% in cattle, pigs and poultry used for meat production,
CC reducing the need for antibiotic feed-additives. Anti-GDF8 vaccines can
CC be used to treat human diseases such as cancer cachexia where muscle
CC atrophy is pronounced and for patients suffering from acute and chronic
CC heart failure
XX
SQ Sequence 109 AA;

Query Match 59.0%; Score 118; DB 4; Length 109;
Best Local Similarity 61.0%; Pred. No. 3e-09;
Matches 25; Conservative 2; Mismatches 8; Indels 6; Gaps 1;

```
Qy          3 WSYGLRPG------SSGPSLFNNFTVSFWLRVPKVSASHLE 37
              | : : |        ||   ||||||||||||||||||||
Db         29 WDWIIAPKRYKANYCSGECEFNNFTVSFWLRVPKVSASHLE 69
```

RESULT 12
AAY92662
ID AAY92662 standard; protein; 703 AA.
XX
AC AAY92662;
XX
DT 10-AUG-2000 (first entry)
XX
DE Mutant murine PSM splice variant construct, mPSM'Y.
XX
KW Prostate specific membrane antigen; splice variant; mutant; vaccination;
KW cytotoxic T-lymphocyte immunity; breast cancer; prostate cancer;
KW cell-associated peptide antigen; foreign epitope.
XX
OS Mus musculus.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Peptide 631. .651
FT /label= P30
XX
PN WO200020027-A2.
XX
PD 13-APR-2000.
XX
PF 05-OCT-1999; 99WO-DK000525.
XX
PR 05-OCT-1998; 98DK-00001261.
PR 20-OCT-1998; 98US-0105011P.
XX
PA (MEBI-) M & E BIOTECH AS.
XX
PI Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
PI Gautam A, Birk P, Karlsson G;
XX
DR WPI; 2000-349917/30.
XX
PT Inducing immune responses to weakly immunogenic, tumor associated peptide
PT antigens for the treatment of breast and prostate cancer.
XX
PS Example 1; Page; 220pp; English.
XX
CC AAY92659-62 are mutant immunogenized murine prostate specific membrane
CC antigen (PSM) constructs, which contain a foreign epitope, P30. The
CC analogues can be used to study whether autotolerance to mouse PSM can be
CC broken in mice by immunisation and/or DNA vaccination against murine PSM
CC using murine PSM analogues. Immunogenic analogues of PSM can be used in
CC the claimed method as an autovaccine to induce a CTL response. The method
CC is used for inducing immune responses against weakly immunogenic cell-
CC associated peptide antigens (PA) such as those associated with cancers
CC (self-proteins), e.g. human PSM, heregulin 2 (Her2) and/or fibroblast
CC growth factor 8b (FGF8b). The method comprises effecting simultaneous
CC presentation by antigen producing cells (APCs) of the animals immune
CC system of: (1) at least 1 CTL (cytotoxic T-lymphocyte) group derived from
CC the PA and/or at least 1 B-cell group derived from the cell-associated PA
CC ; and (2) at least 1 first T helper cell group which is foreign to the
CC animal. Analogues of human PSM, human Her2 and human/murine FGF8b
CC comprising a substantial part of all known and predicted CTL and B-cell
CC epitopes of the respective PA and including at least one foreign T helper
CC epitope are also claimed. The method is used to treat prostate,
CC prostate/breast or breast cancer when the PA is human PSM, FGF8b and
CC Her2, respectively. Note: This sequence was constructed from the murine
CC PSM splice variant (AAY92624), which appears on pages 210-213 of the
CC specification
XX
SQ Sequence 703 AA;

Query Match 59.0%; Score 118; DB 3; Length 703;
Best Local Similarity 82.1%; Pred. No. 2.6e-08;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy         10 GSSGPSLFNNFTVSFWLRVPKVSASHLE 37
              |  |   ||||||||||||||||||||||
Db        624 GLPGRPFFNNFTVSFWLRVPKVSASHLE 651
```

RESULT 13
AAY92660
ID AAY92660 standard; protein; 761 AA.
XX
AC AAY92660;
XX
DT 10-AUG-2000 (first entry)
XX
DE Mutant murine prostate specific membrane antigen construct, mPSMY.
XX
KW Prostate specific membrane antigen; immunogenized construct; mutant;
KW vaccination; cytotoxic T-lymphocyte immunity; breast cancer;
KW prostate cancer; cell-associated peptide antigen; foreign epitope.
XX
OS Mus musculus.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Peptide 689. .709
FT /label= P30
XX
PN WO200020027-A2.
XX
PD 13-APR-2000.
XX
PF 05-OCT-1999; 99WO-DK000525.
XX
PR 05-OCT-1998; 98DK-00001261.
PR 20-OCT-1998; 98US-0105011P.

XX
PA (MEBI-) M & E BIOTECH AS.
XX
PI Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
PI Gautam A, Birk P, Karlsson G;
XX
DR WPI; 2000-349917/30.
XX
PT Inducing immune responses to weakly immunogenic, tumor associated peptide
PT antigens for the treatment of breast and prostate cancer.
XX
PS Example 1; Page; 220pp; English.
XX
CC AAY92659-62 are mutant immunogenized murine prostate specific membrane
CC antigen (PSM) constructs, which contain a foreign epitope, P30. The
CC analogues can be used to study whether autotolerance to mouse PSM can be
CC broken in mice by immunisation and/or DNA vaccination against murine PSM
CC using murine PSM analogues. Immunogenic analogues of PSM can be used in
CC the claimed method as an autovaccine to induce a CTL response. The method
CC is used for inducing immune responses against weakly immunogenic cell-
CC associated peptide antigens (PA) such as those associated with cancers
CC (self-proteins), e.g. human PSM, heregulin 2 (Her2) and/or fibroblast
CC growth factor 8b (FGF8b). The method comprises effecting simultaneous
CC presentation by antigen producing cells (APCs) of the animals immune
CC system of: (1) at least 1 CTL (cytotoxic T-lymphocyte) group derived from
CC the PA and/or at least 1 B-cell group derived from the cell-associated PA
CC ; and (2) at least 1 first T helper cell group which is foreign to the
CC animal. Analogues of human PSM, human Her2 and human/murine FGF8b
CC comprising a substantial part of all known and predicted CTL and B-cell
CC epitopes of the respective PA and including at least one foreign T helper
CC epitope are also claimed. The method is used to treat prostate,
CC prostate/breast or breast cancer when the PA is human PSM, FGF8b and
CC Her2, respectively. Note: This sequence was constructed from the wild
CC type murine PSM (AAY92623), which appears on pages 204-206 of the
CC specification
XX
SQ Sequence 761 AA;

Query Match 59.0%; Score 118; DB 3; Length 761;
Best Local Similarity 82.1%; Pred. No. 2.9e-08;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 10 GSSGPSLFNNFTVSFWLRVPKVSASHLE 37
| | |||||||||||||||||||||||||||
Db 682 GLPGRPFFNNFTVSFWLRVPKVSASHLE 709

RESULT 14
AAB45496
ID AAB45496 standard; protein; 124 AA.
XX
AC AAB45496;
XX
DT 26-FEB-2001 (first entry)
XX
DE Modified human interleukin-5 SEQ ID NO: 8.
XX
KW Asthma; IL-5; interleukin-5; allergy; cytokine; helminthic infection;
KW cancer; eosinophilia; vaccine; allergic rhinitis.
XX
OS Homo sapiens.
OS Clostridium tetani.
XX
PN WO200065058-A1.
XX
PD 02-NOV-2000.
XX
PF 19-APR-2000; 2000WO-DK000205.
XX
PR 23-APR-1999; 99DK-00000552.
PR 06-MAY-1999; 99US-0132811P.
XX
PA (MEBI-) M & E BIOTECH AS.
XX
PI Klysner S;
XX
DR WPI; 2000-672791/65.
XX
PT Down-regulating interleukin 5 (IL-5) activity in humans by administering
PT IL-5 and/or an IL-5 analogue, useful in the treatment, prophylaxis or
PT amelioration of asthma or other chronic allergic conditions.
XX
PS Example 7; Page 124; 172pp; English.
XX
CC The present invention is concerned with methods of treating asthma,
CC eosinophilia, allergic rhinitis and other allergic diseases. These
CC involve the use of interleukin-5 (IL-5) analogues and modified IL-5
CC proteins and their coding sequences to down-regulate IL-5 activity and
CC thus reduce eosinophil numbers. The allergic diseases may be treated
CC using autovaccines, nucleic acid vaccines or live vaccines. In addition,
CC it is possible that they may be used in the treatment of cancer and
CC helminthic infections
XX
SQ Sequence 124 AA;

Query Match 58.5%; Score 117; DB 3; Length 124;
Best Local Similarity 95.7%; Pred. No. 4.9e-09;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 15 SLFNNFTVSFWLRVPKVSASHLE 37
:|||||||||||||||||||||||
Db 30 TLFNNFTVSFWLRVPKVSASHLE 52

RESULT 15
AAB45515
ID AAB45515 standard; protein; 124 AA.
XX
AC AAB45515;
XX
DT 26-FEB-2001 (first entry)
XX
DE Modified human interleukin-5 SEQ ID NO: 30.
XX
KW Asthma; IL-5; interleukin-5; allergy; cytokine; helminthic infection;
KW cancer; eosinophilia; vaccine; allergic rhinitis.
XX
OS Homo sapiens.
OS Clostridium tetani.
XX
PN WO200065058-A1.
XX
PD 02-NOV-2000.
XX
PF 19-APR-2000; 2000WO-DK000205.
XX
PR 23-APR-1999; 99DK-00000552.
PR 06-MAY-1999; 99US-0132811P.
XX
PA (MEBI-) M & E BIOTECH AS.
XX
PI Klysner S;
XX
DR WPI; 2000-672791/65.
DR N-PSDB; AAC68868.
XX
PT Down-regulating interleukin 5 (IL-5) activity in humans by administering
PT IL-5 and/or an IL-5 analogue, useful in the treatment, prophylaxis or
PT amelioration of asthma or other chronic allergic conditions.
XX
PS Example 7; Page 141; 172pp; English.
XX
CC The present invention is concerned with methods of treating asthma,
CC eosinophilia, allergic rhinitis and other allergic diseases. These
CC involve the use of interleukin-5 (IL-5) analogues and modified IL-5
CC proteins and their coding sequences to down-regulate IL-5 activity and

CC thus reduce eosinophil numbers. The allergic diseases may be treated
CC using autovaccines, nucleic acid vaccines or live vaccines. In addition,
CC it is possible that they may be used in the treatment of cancer and
CC helminthic infections
XX
SQ Sequence 124 AA;

Query Match 58.5%; Score 117; DB 3; Length 124;
Best Local Similarity 95.7%; Pred. No. 4.9e-09;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy         15 SLFNNFTVSFWLRVPKVSASHLE 37
              :|||||||||||||||||||||||
Db         30 TLFNNFTVSFWLRVPKVSASHLE 52
```

Search completed: March 10, 2004, 09:12:11
Job time : 55.716 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:53 ; Search time 6.76654 Seconds
(without alignments)
284.724 Million cell updates/sec

Title: US-09-848-834A-14
Perfect score: 200
Sequence: 1 XHWSYGLRPGSSGPSLFNNFTVSFWLRVPKVSASHLE 37

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	114	57.0	1314	1	TETX_CLOTE	P04958 clostridium
2	67	33.5	1296	1	BXG_CLOBO	Q60393 clostridium
3	65	32.5	1295	1	BXA1_CLOBO	P10845 clostridium
4	63	31.5	1274	1	BXF_CLOBO	P30996 clostridium
5	63	31.5	1290	1	BXB_CLOBO	P10844 clostridium
6	62	31.0	1051	1	VP2_AHSV6	O71024 african hor
7	62	31.0	1295	1	BXA2_CLOBO	Q45894 clostridium
8	60	30.0	92	1	GON1_TUPGB	Q95335 tupaia glis
9	59	29.5	67	1	GON1_MACMU	P55247 macaca mula
10	59	29.5	92	1	GON1_HUMAN	P01148 homo sapien
11	58	29.0	61	1	GON1_SHEEP	Q28588 ovis aries
12	58	29.0	63	1	GON1_MESAU	O09163 mesocricetu
13	58	29.0	89	1	GON1_XENLA	P45656 xenopus lae
14	58	29.0	90	1	GON1_MOUSE	P13562 mus musculu
15	58	29.0	90	1	GON1_RANCA	Q90y63 rana catesb
16	58	29.0	91	1	GON1_PIG	P49921 sus scrofa
17	58	29.0	92	1	GON1_RAT	P07490 rattus norv
18	57.5	28.7	352	1	COA2_SV40	P03093 simian viru
19	56.5	28.2	94	1	GON1_HAPBU	P51918 haplochromi
20	56	28.0	92	1	GON1_CHICK	P37042 gallus gall
21	56	28.0	1250	1	BXE_CLOBO	Q00496 clostridium
22	56	28.0	1250	1	BXE_CLOBU	P30995 clostridium
23	56	28.0	1290	1	BXC1_CLOBO	P18640 clostridium
24	54.5	27.3	91	1	GON1_ORYLA	Q9dgc8 o progonado
25	54.5	27.3	464	1	VNSS_TSWV1	P26002 tomato spot
26	54	27.0	10	1	GON1_ALLMI	P37041 alligator m
27	53	26.5	99	1	GON1_DICLA	Q9ia10 dicentrarch
28	53	26.5	449	1	VNSS_INSVN	Q01811 impatiens n
29	53	26.5	1196	1	BXCN_CLOBO	P46081 clostridium
30	52	26.0	95	1	GON1_MORSA	O73812 morone saxa
31	52	26.0	95	1	GON1_PAGMA	P70074 pagrus majo
32	52	26.0	95	1	GON1_SPAAU	P51919 sparus aura
33	52	26.0	467	1	VNSS_TSWVL	P26003 tomato spot
34	51.5	25.8	1539	1	SMCY_HUMAN	Q9by66 homo sapien
35	51.5	25.8	1548	1	SMCY_MOUSE	Q62240 mus musculu
36	51.5	25.8	1554	1	SMCX_MOUSE	P41230 mus musculu
37	51.5	25.8	1560	1	SMCX_HUMAN	P41229 homo sapien
38	51	25.5	227	1	RHON_HUMAN	P52198 homo sapien
39	51	25.5	227	1	RHON_MOUSE	Q9qym5 mus musculu
40	51	25.5	956	1	CLS3_HUMAN	Q9bqt9 homo sapien
41	50.5	25.2	1276	1	BXD_CLOBO	P19321 clostridium
42	50	25.0	90	1	GON8_RANDY	Q9iau2 rana dybows
43	50	25.0	144	1	CNI_DROME	P49858 drosophila
44	50	25.0	183	1	RRF_BIFLO	Q8g483 bifidobacte
45	50	25.0	585	1	YH70_SYNY3	P73627 synechocyst

ALIGNMENTS

RESULT 1
TETX_CLOTE
ID TETX_CLOTE STANDARD; PRT; 1314 AA.
AC P04958;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Tetanus toxin precursor (EC 3.4.24.68) (Tentoxylysin) [Contains:
DE Tetanus toxin light chain (Tetanus toxin chain L); Tetanus toxin heavy
DE chain (Tetanus toxin chain H)].
GN TETX OR CTP60.
OS Clostridium tetani.
OG Plasmid pE88, and Plasmid 75 Kbp.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1513;
RN [1]
RP SEQUENCE FROM N.A.
RC PLASMID=75 Kbp;
RX MEDLINE=87053814; PubMed=3536478;
RA Eisel U., Jarausch W., Goretzki K., Henschen A., Engels J.,
RA Weller U., Hudel M., Habermann E., Niemann H.;
RT "Tetanus toxin: primary structure, expression in E. coli, and
RT homology with botulinum toxins.";
RL EMBO J. 5:2495-2502(1986).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CN3911; PLASMID=75 Kbp;
RX MEDLINE=87040747; PubMed=3774547;
RA Fairweather N.F., Lyness V.A.;
RT "The complete nucleotide sequence of tetanus toxin.";
RL Nucleic Acids Res. 14:7809-7812(1986).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Massachusetts / E88; PLASMID=pE88;
RX MEDLINE=22457253; PubMed=12552129;
RA Brueggemann H., Baeumer S., Fricke W.F., Wiezer A., Liesegang H.,
RA Decker I., Herzberg C., Martinez-Arias R., Merkl R., Henne A.,
RA Gottschalk G.;
RT "The genome sequence of Clostridium tetani, the causative agent of
RT tetanus disease.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:1316-1321(2003).
RN [4]
RP SEQUENCE OF 742-1314 FROM N.A.
RC PLASMID=75 Kbp;
RX MEDLINE=86085672; PubMed=3510187;
RA Fairweather N.F., Lyness V.A., Pickard D.J., Allen G., Thomson R.O.;
RT "Cloning, nucleotide sequencing, and expression of tetanus toxin
RT fragment C in Escherichia coli.";
RL J. Bacteriol. 165:21-27(1986).
RN [5]
RP PARTIAL SEQUENCE, AND DISULFIDE BONDS.
RX MEDLINE=90201034; PubMed=2108021;
RA Krieglstein K., Henschen A., Weller U., Habermann E.;
RT "Arrangement of disulfide bridges and positions of sulfhydryl groups
RT in tetanus toxin.";

```
RL   Eur. J. Biochem. 188:39-45(1990).
RN   [6]
RP   PARTIAL SEQUENCE.
RX   MEDLINE=92037649; PubMed=1935979;
RA   Krieglstein K.G., Henschen A.H., Weller U., Habermann E.;
RT   "Limited proteolysis of tetanus toxin. Relation to activity and
RT   identification of cleavage sites.";
RL   Eur. J. Biochem. 202:41-51(1991).
RN   [7]
RP   IDENTIFICATION AS ZINC-PROTEASE.
RX   MEDLINE=93010948; PubMed=1396558;
RA   Schiavo G., Poulain B., Rossetto O., Benfenati F., Tauc L.,
RA   Montecucco C.;
RT   "Tetanus toxin is a zinc protein and its inhibition of
RT   neurotransmitter release and protease activity depend on zinc.";
RL   EMBO J. 11:3577-3583(1992).
RN   [8]
RP   IDENTIFICATION OF SUBSTRATE.
RX   MEDLINE=93063293; PubMed=1331807;
RA   Schiavo G., Benfenati F., Poulain B., Rossetto O., de Laureto P.P.,
RA   Dasgupta B.R., Montecucco C.;
RT   "Tetanus and botulinum-B neurotoxins block neurotransmitter release
RT   by proteolytic cleavage of synaptobrevin.";
RL   Nature 359:832-835(1992).
RN   [9]
RP   X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 874-1314.
RX   MEDLINE=97475217; PubMed=9334741;
RA   Umland T.C., Wingert L.M., Swaminathan S., Furey W.F., Schmidt J.J.,
RA   Sax M.;
RT   "Structure of the receptor binding fragment HC of tetanus
RT   neurotoxin.";
RL   Nat. Struct. Biol. 4:788-792(1997).
CC   -!- FUNCTION: TETANUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC       RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC       AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC       WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC       INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC       ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 76-GLN-|-PHE-77
CC       BOND OF SYNAPTOBREVIN-2.
CC   -!- CATALYTIC ACTIVITY: Hydrolysis of 76-Gln-|-Phe-77 bond in
CC       synaptobrevin 2.
CC   -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC   -!- SUBUNIT: THE PRECURSOR POLYPEPTIDE IS SUBSEQUENTLY CLEAVED TO
CC       YIELD SUBCHAINS L AND H. THESE REMAIN LINKED BY A DISULFIDE BRIDGE
CC       AND ARE NON-TOXIC AFTER SEPARATION.
CC   -!- MISCELLANEOUS: THE C-TERMINAL OF THE HEAVY CHAIN BINDS TO
CC       GANGLIOSIDE RECEPTORS.
CC   -!- SIMILARITY: Belongs to peptidase family M27.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X04436; CAA28033.1; -.
DR   EMBL; X06214; CAA29564.1; -.
DR   EMBL; AF528097; AAO37454.1; -.
DR   EMBL; M12739; AAA23282.1; -.
DR   PIR; A25689; BTCLTN.
DR   PDB; 1AF9; 29-APR-98.
DR   PDB; 1A8D; 14-OCT-98.
DR   PDB; 1D0H; 27-MAR-00.
DR   PDB; 1DFQ; 24-MAR-00.
DR   PDB; 1DIW; 24-MAR-00.
DR   PDB; 1DLL; 24-MAR-00.
DR   PDB; 1FV3; 05-SEP-01.
DR   MEROPS; M27.001; -.
DR   InterPro; IPR008985; ConA_like_lec_gl.
DR   InterPro; IPR002160; Kunitz_legume.
DR   InterPro; IPR006025; Pept_M_Zn_BS.
DR   InterPro; IPR000395; Peptidase_M27.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
KW   Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc; Plasmid;
KW   3D-structure; Complete proteome.
FT   INIT_MET      0      0
FT   CHAIN         1    456       TETANUS TOXIN LIGHT CHAIN.
FT   CHAIN       457   1314       TETANUS TOXIN HEAVY CHAIN.
FT   METAL       232    232       ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT_SITE    233    233       BY SIMILARITY.
FT   METAL       236    236       ZINC (CATALYTIC) (BY SIMILARITY).
FT   TRANSMEM    226    246       POTENTIAL.
FT   TRANSMEM    669    689       POTENTIAL.
FT   DISULFID    438    466       INTERCHAIN.
FT   DISULFID   1076   1092
FT   HELIX       876    882
FT   TURN        883    883
FT   STRAND      884    891
FT   TURN        892    893
FT   STRAND      894    897
FT   STRAND      904    907
FT   TURN        909    910
FT   STRAND      912    915
FT   STRAND      920    925
FT   TURN        928    929
FT   STRAND      932    935
FT   HELIX       938    940
FT   TURN        941    946
FT   STRAND      949    956
FT   HELIX       962    968
FT   TURN        969    970
FT   STRAND      972    977
FT   STRAND      980    981
FT   HELIX       983    985
FT   STRAND      987    995
FT   TURN        996    997
FT   STRAND      998   1004
FT   TURN       1006   1007
FT   STRAND     1010   1016
FT   STRAND     1020   1020
FT   TURN       1021   1022
FT   STRAND     1031   1037
FT   TURN       1039   1040
FT   STRAND     1042   1047
FT   TURN       1048   1049
FT   STRAND     1050   1056
FT   TURN       1058   1059
FT   STRAND     1068   1074
FT   TURN       1079   1080
FT   STRAND     1082   1091
FT   HELIX      1097   1105
FT   TURN       1106   1107
FT   STRAND     1112   1112
FT   STRAND     1114   1114
FT   TURN       1116   1117
FT   STRAND     1120   1120
FT   STRAND     1122   1122
FT   TURN       1123   1124
FT   STRAND     1127   1131
FT   HELIX      1132   1134
FT   TURN       1135   1136
FT   STRAND     1137   1141
FT   TURN       1144   1145
FT   STRAND     1148   1152
FT   STRAND     1155   1158
FT   TURN       1159   1162
FT   STRAND     1163   1166
FT   STRAND     1173   1178
FT   TURN       1184   1185
FT   STRAND     1188   1188
FT   STRAND     1190   1190
```

```
  Query Match             57.0%;  Score 114;  DB 1;  Length 1314;
  Best Local Similarity   95.5%;  Pred. No. 5.6e-08;
  Matches   21;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         16 LFNNFTVSFWLRVPKVSASHLE 37
              :||||||||||||||||||||||
Db        945 MFNNFTVSFWLRVPKVSASHLE 966


RESULT 2
BXG_CLOBO
ID   BXG_CLOBO      STANDARD;      PRT;  1296 AA.
AC   Q60393;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Botulinum neurotoxin type G precursor (EC 3.4.24.69) (BoNT/G)
DE   (Bontoxilysin G).
GN   BOTG.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=113 / 30;
RX   MEDLINE=94092745; PubMed=8268233;
RA   Campbell K., Collins M.D., East A.K.;
RT   "Nucleotide sequence of the gene coding for Clostridium botulinum
RT   (Clostridium argentinense) type G neurotoxin: genealogical comparison
RT   with other clostridial neurotoxins.";
RL   Biochim. Biophys. Acta 1216:487-491(1993).
CC   -!- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC       RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC       AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC       WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC       INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC       ENDOPEPTIDASE.
CC   -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC       neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
CC       detected action on small molecule substrates.
CC   -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC   -!- SUBUNIT: Disulfide-linked heterodimer of a light chain (L) and a
CC       heavy chain (H). The light chain has the pharmacological activity,
CC       while the N- and C-terminal of the heavy chain mediate channel
CC       formation and toxin binding, respectively.
CC   -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC   -!- MISCELLANEOUS: There are seven antigenically distinct forms of
CC       botulinum neurotoxin: Types A, B, C1, D, E, F, and G.
CC   -!- SIMILARITY: Belongs to peptidase family M27.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X74162; CAA52275.1; -.
DR   HSSP; P10845; 3BTA.
DR   MEROPS; M27.002; -.
DR   InterPro; IPR008985; ConA_like_lec_gl.
DR   InterPro; IPR002160; Kunitz_legume.
DR   InterPro; IPR006025; Pept_M_Zn_BS.
DR   InterPro; IPR000395; Peptidase_M27.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
KW   Neurotoxin; Hydrolase; Metalloprotease; Zinc.
FT   INIT_MET      0      0       BY SIMILARITY.
FT   CHAIN         1    441       BOTULINUM NEUROTOXIN G, LIGHT-CHAIN.
FT   CHAIN       442   1296       BOTULINUM NEUROTOXIN G, HEAVY-CHAIN.
FT   METAL       229    229       ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT_SITE    230    230       BY SIMILARITY.
FT   METAL       233    233       ZINC (CATALYTIC) (BY SIMILARITY).
FT   DISULFID    435    449       INTERCHAIN (PROBABLE).
SQ   SEQUENCE   1296 AA;  149013 MW;  DC8E47E15F665C31 CRC64;

  Query Match             33.5%;  Score 67;  DB 1;  Length 1296;
  Best Local Similarity   39.1%;  Pred. No. 0.26;
  Matches    9;  Conservative   10;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         15 SLFNNFTVSFWLRVPKVSASHLE 37
              |:|:||:::||:| || : : ::
Db        927 SMFDNFSINFWVRTPKYNNNDIQ 949


RESULT 3
BXA1_CLOBO
ID   BXA1_CLOBO     STANDARD;      PRT;  1295 AA.
AC   P10845; P01561; P18639;
DT   01-JUL-1989 (Rel. 11, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Botulinum neurotoxin type A precursor (EC 3.4.24.69) (BoNT/A)
DE   (Bontoxilysin A) (BOTOX) [Contains: Botulinum neurotoxin A, light-
DE   chain; Botulinum neurotoxin A, heavy-chain].
GN   BOTA OR BNA OR ATX.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NCTC 2916;
RX   MEDLINE=90235864; PubMed=2185020;
RA   Thompson D.E., Brehm J.K., Oultram J.D., Swinfield T.-J.,
RA   Shone C.C., Atkinson T., Melling J., Minton N.P.;
RT   "The complete amino acid sequence of the Clostridium botulinum type A
RT   neurotoxin, deduced by nucleotide sequence analysis of the encoding
RT   gene.";
RL   Eur. J. Biochem. 189:73-81(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=62A;
RX   MEDLINE=90264400; PubMed=2160960;
RA   Binz B., Kuarzono H., Wille M., Frevent J., Wernars K., Niemann H.;
RT   "The complete sequence of botulinum neurotoxin type A and comparison
RT   with other clostridial neurotoxins.";
RL   J. Biol. Chem. 265:9153-9158(1990).
RN   [3]
RP   SEQUENCE OF 1-65 FROM N.A.
RC   STRAIN=62A;
RX   MEDLINE=97016817; PubMed=8863443;
RA   East A.K., Bhandari M., Stacey J.M., Campbell K.D., Collins M.D.;
RT   "Organization and phylogenetic interrelationships of genes encoding
RT   components of the botulinum toxin complex in proteolytic Clostridium
RT   botulinum types A, B, and F: evidence of chimeric sequences in the
RT   gene encoding the nontoxic nonhemagglutinin component.";
RL   Int. J. Syst. Bacteriol. 46:1105-1112(1996).
RN   [4]
RP   SEQUENCE OF 1-34 FROM N.A.
RC   STRAIN=Hall;
RX   MEDLINE=89350959; PubMed=2669749;
RA   Betley M.J., Somers E., Dasgupta B.R.;
RT   "Characterization of botulinum type A neurotoxin gene: delineation of
RT   the N-terminal encoding region.";
RL   Biochem. Biophys. Res. Commun. 162:1388-1395(1989).
RN   [5]
RP   SEQUENCE OF 1-18 FROM N.A.
RC   STRAIN=Type A NIH;
RX   MEDLINE=96096783; PubMed=8521962;
RA   Fujita R., Fujinaga Y., Inoue K., Nakajima H., Kumon H., Oguma K.;
```

```
RT   "Molecular characterization of two forms of nontoxic-nonhemagglutinin
RT   components of Clostridium botulinum type A progenitor toxins.";
RL   FEBS Lett. 376:41-44(1995).
RN   [6]
RP   SEQUENCE OF 1-16.
RX   MEDLINE=84178501; PubMed=6370252;
RA   Schmidt J.J., Sartymoorthy V., Dasgupta B.R.;
RT   "Partial amino acid sequence of the heavy and light chains of
RT   botulinum neurotoxin type A.";
RL   Biochem. Biophys. Res. Commun. 119:900-904(1984).
RN   [7]
RP   SEQUENCE OF 1-46.
RA   Dasgupta B.R., Foley J., Niece R.;
RT   "Partial sequence of the light chain of botulinum neurotoxin type A.";
RL   Biochemistry 26:4162-4162(1987).
RN   [8]
RP   SEQUENCE OF 1-5 AND 444-456.
RX   MEDLINE=91120847; PubMed=2126206;
RA   Dasgupta B.R., Dekleva M.L.;
RT   "Botulinum neurotoxin type A: sequence of amino acids at the
RT   N-terminus and around the.nicking site.";
RL   Biochimie 72:661-664(1990).
RN   [9]
RP   SEQUENCE OF 448-464 AND 872-895.
RX   MEDLINE=89024662; PubMed=3178218;
RA   Sathymoorthy V., Dasgupta B.R., Foley J., Niece R.L.;
RT   "Botulinum neurotoxin type A: cleavage of the heavy chain into two
RT   halves and their partial sequences.";
RL   Arch. Biochem. Biophys. 266:142-151(1988).
RN   [10]
RP   SEQUENCE OF 448-482.
RX   MEDLINE=85285016; PubMed=3896784;
RA   Shone C.C., Hambleton P., Melling J.;
RT   "Inactivation of Clostridium botulinum type A neurotoxin by trypsin
RT   and purification of two tryptic fragments. Proteolytic action near
RT   the COOH-terminus of the heavy subunit destroys toxin-binding
RT   activity.";
RL   Eur. J. Biochem. 151:75-82(1985).
RN   [11]
RP   IDENTIFICATION OF SUBSTRATE.
RX   MEDLINE=94063091; PubMed=8243676;
RA   Schiavo G., Santtuci A., Dasgupta B.R., Mehta P.P., Jontes J.,
RA   Benfenati F., Wilson M.C., Montecucco C.;
RT   "Botulinum neurotoxins serotypes A and E cleave SNAP-25 at distinct
RT   COOH-terminal peptide bonds.";
RL   FEBS Lett. 335:99-103(1993).
RN   [12]
RP   IDENTIFICATION OF SUBSTRATE.
RX   MEDLINE=94124495; PubMed=8294407;
RA   Binz T., Blasi J., Yamasaki S., Baumeister A., Link E., Suedhof T.C.,
RA   Jahn R., Niemann H.;
RT   "Proteolysis of SNAP-25 by types E and A botulinal neurotoxins.";
RL   J. Biol. Chem. 269:1617-1620(1994).
RN   [13]
RP   MUTAGENESIS OF GLU-261; PHE-265 AND TYR-365.
RX   MEDLINE=21556941; PubMed=11700044;
RA   Rigoni M., Caccin P., Johnson E.A., Montecucco C., Rossetto O.;
RT   "Site-directed mutagenesis identifies active-site residues of the
RT   light chain of botulinum neurotoxin type a.";
RL   Biochem. Biophys. Res. Commun. 288:1231-1237(2001).
RN   [14]
RP   X-RAY CRYSTALLOGRAPHY (3.3 ANGSTROMS).
RX   MEDLINE=98455071; PubMed=9783750;
RA   Lacy D.B., Tepp W., Cohen A.C., Dasgupta B.R., Stevens R.C.;
RT   "Crystal structure of botulinum neurotoxin type A and implications
RT   for toxicity.";
RL   Nat. Struct. Biol. 5:898-902(1998).
CC   -!- FUNCTION: Inhibits acetylcholine release. The botulinum toxin
CC       binds with high affinity to peripheral neuronal presynaptic
CC       membrane, is then internalized by receptor-mediated endocytosis.
CC       The C-terminus of the heavy chain (H) is responsible for the
CC       adherence of the toxin to the cell surface while the N-terminus
CC       mediates transport of the light chain from the endocytic vesicle
CC       to the cytosol. After translocation, the light chain (L)
CC       hydrolyzes the 197-Gln-|-Arg-198 bond in SNAP-25, thereby blocking
CC       neurotransmitter release. Inhibition of acetylcholine release
CC       results in flaccid paralysis, with frequent heart or respiratory
CC       failure.
CC   -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC       neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
CC       detected action on small molecule substrates.
CC   -!- COFACTOR: Binds 1 zinc ion per subunit.
CC   -!- SUBUNIT: Disulfide-linked heterodimer of a light chain (L) and a
CC       heavy chain (H).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- PHARMACEUTICAL: Available under the name BOTOX (Allergan) for
CC       the treatment of strabismus and blepharospasm associated with
CC       dystonia and cervical dystonia. Also used for the treatment of
CC       hemifacial spasm and a number of other neurological disorders
CC       characterized by abnormal muscle contraction.
CC   -!- MISCELLANEOUS: There are seven antigenically distinct forms of
CC       botulinum neurotoxin: Types A, B, C1, D, E, F, and G.
CC   -!- SIMILARITY: Belongs to peptidase family M27.
CC   -!- DATABASE: NAME=BOTOX product information Web site;
CC       WWW="http://www.botox.com/index.jsp?hp&productinfo".
CC   -!- DATABASE: NAME=Protein Spotlight;
CC       NOTE=Issue 19 of February 2002;
CC       WWW="http://www.expasy.org/spotlight/articles/sptlt019.html".
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; X52066; CAA36289.1; -.
DR   EMBL; M30196; AAA23262.1; -.
DR   EMBL; X92973; CAA63551.1; -.
DR   EMBL; D67030; BAA11051.1; -.
DR   EMBL; M27892; AAA23269.1; -.
DR   PIR; A35294; BTCLAB.
DR   PDB; 3BTA; 01-OCT-99.
DR   MEROPS; M27.002; -.
DR   InterPro; IPR008985; ConA_like_lec_gl.
DR   InterPro; IPR002160; Kunitz_legume.
DR   InterPro; IPR006025; Pept_M_Zn_BS.
DR   InterPro; IPR000395; Peptidase_M27.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
KW   Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc;
KW   Pharmaceutical; 3D-structure.
FT   INIT_MET      0      0
FT   CHAIN         1    447       BOTULINUM NEUROTOXIN A, LIGHT-CHAIN.
FT   CHAIN       448   1295       BOTULINUM NEUROTOXIN A, HEAVY-CHAIN.
FT   METAL       222    222       ZINC (CATALYTIC).
FT   ACT_SITE    223    223
FT   METAL       226    226       ZINC (CATALYTIC).
FT   METAL       261    261       ZINC (CATALYTIC).
FT   DISULFID    429    453       INTERCHAIN.
FT   DISULFID   1234   1279
FT   TRANSMEM    626    646       POTENTIAL.
FT   TRANSMEM    655    675       POTENTIAL.
FT   VARIANT      26     26       V -> A.
FT   MUTAGEN     261    261       E->A: DRASTIC DECREASE IN ENZYMATIC
FT                                ACTIVITY.
FT   MUTAGEN     265    265       F->A: DECREASES ENZYMATIC ACTIVITY.
FT   MUTAGEN     365    365       Y->A: DECREASES ENZYMATIC ACTIVITY.
FT   CONFLICT      1      1       P -> Q (IN REF. 1).
FT   CONFLICT    479    479       E -> P (IN REF. 9).
FT   CONFLICT    875    875       T -> L (IN REF. 8).
FT   CONFLICT    891    891       S -> K (IN REF. 8).
SQ   SEQUENCE   1295 AA;  149322 MW;  858342F754862579 CRC64;
```

```
  Query Match             32.5%;  Score 65;  DB 1;  Length 1295;
  Best Local Similarity   56.2%;  Pred. No. 0.51;
  Matches    9;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         15 SLFNNFTVSFWLRVPK 30
              |:: ||: |||:|:||
Db        935 SMYENFSTSFWIRIPK 950


RESULT 4
BXF_CLOBO
ID   BXF_CLOBO      STANDARD;      PRT;  1274 AA.
AC   P30996;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Botulinum neurotoxin type F precursor (EC 3.4.24.69) (BoNT/F)
DE   (Bontoxilysin F).
GN   BOTF.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 23387;
RX   MEDLINE=93012902; PubMed=1398040;
RA   East A.K., Richardson P.T., Allaway D., Collins M.D.,
RA   Roberts T.A., Thompson D.E.;
RT   "Sequence of the gene encoding type F neurotoxin of Clostridium
RT   botulinum.";
RL   FEMS Microbiol. Lett. 75:225-230(1992).
RN   [2]
RP   SEQUENCE OF 1-64 FROM N.A.
RC   STRAIN=Hobbs FT10;
RX   MEDLINE=94297488; PubMed=7764998;
RA   East A.K., Collins M.D.;
RT   "Conserved structure of genes encoding components of botulinum
RT   neurotoxin complex M and the sequence of the gene coding for the
RT   nontoxic component in nonproteolytic Clostridium botulinum type F.";
RL   Curr. Microbiol. 29:69-77(1994).
RN   [3]
RP   SEQUENCE OF 634-1002 FROM N.A.
RX   MEDLINE=94013372; PubMed=8408542;
RA   Campbell K., East A.K., Collins M.D.;
RT   "Gene probes for identification of the botulinal neurotoxin gene and
RT   specific identification of neurotoxin types B, E, and F.";
RL   J. Clin. Microbiol. 31:2255-2262(1993).
RN   [4]
RP   IDENTIFICATION OF SUBSTRATE.
RX   MEDLINE=94230352; PubMed=8175689;
RA   Yamasaki S., Baumeister A., Binz T., Blasi J., Link E., Cornille F.,
RA   Roques B., Fykse E.M., Suedhof T.C., Jahn R., Niemann H.;
RT   "Cleavage of members of the synaptobrevin/VAMP family by types D and
RT   F botulinal neurotoxins and tetanus toxin.";
RL   J. Biol. Chem. 269:12764-12772(1994).
CC   -!- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC       RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC       AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC       WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC       INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC       ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 58-GLN-|-LYS-59
CC       BOND OF SYNAPTOBREVINS-1 AND -2.
CC   -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC       neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
CC       detected action on small molecule substrates.
CC   -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC   -!- SUBUNIT: Disulfide-linked heterodimer of a light chain (L) and a
CC       heavy chain (H). The light chain has the pharmacological activity,
CC       while the N- and C-terminal of the heavy chain mediate channel
CC       formation and toxin binding, respectively.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- MISCELLANEOUS: There are seven antigenically distinct forms of
CC       botulinum neurotoxin: Types A, B, C1, D, E, F, and G.
CC   -!- SIMILARITY: Belongs to peptidase family M27.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M92906; AAA23263.1; -.
DR   EMBL; S73676; AAC60475.1; -.
DR   EMBL; X70820; CAA50151.1; -.
DR   EMBL; X70816; CAA50147.1; -.
DR   PIR; I40813; I40813.
DR   PIR; S48109; S48109.
DR   HSSP; P10845; 3BTA.
DR   MEROPS; M27.002; -.
DR   InterPro; IPR008985; ConA_like_lec_gl.
DR   InterPro; IPR002160; Kunitz_legume.
DR   InterPro; IPR006025; Pept_M_Zn_BS.
DR   InterPro; IPR000395; Peptidase_M27.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
KW   Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
FT   CHAIN         1    436       BOTULINUM NEUROTOXIN F, LIGHT-CHAIN.
FT   CHAIN       437   1274       BOTULINUM NEUROTOXIN F, HEAVY-CHAIN.
FT   METAL       227    227       ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT_SITE    228    228       BY SIMILARITY.
FT   METAL       231    231       ZINC (CATALYTIC) (BY SIMILARITY).
FT   DISULFID    429    445       INTERCHAIN (PROBABLE).
SQ   SEQUENCE   1274 AA;  146709 MW;  5B99756A7438B921 CRC64;

  Query Match             31.5%;  Score 63;  DB 1;  Length 1274;
  Best Local Similarity   56.2%;  Pred. No. 0.96;
  Matches    9;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         15 SLFNNFTVSFWLRVPK 30
              | : ||::|||:|:||
Db        928 SRYQNFSISFWVRIPK 943


RESULT 5
BXB_CLOBO
ID   BXB_CLOBO      STANDARD;      PRT;  1290 AA.
AC   P10844; P10843;
DT   01-JUL-1989 (Rel. 11, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Botulinum neurotoxin type B precursor (EC 3.4.24.69) (BoNT/B)
DE   (Bontoxilysin B).
GN   BOTB.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92384550; PubMed=1514783;
RA   Whelan S.M., Elmore M.J., Bodsworth N.J., Brehm J.K., Atkinson T.,
RA   Minton N.P.;
RT   "Molecular cloning of the Clostridium botulinum structural gene
RT   encoding the type B neurotoxin and determination of its entire
RT   nucleotide sequence.";
RL   Appl. Environ. Microbiol. 58:2345-2354(1992).
RN   [2]
RP   SEQUENCE OF 35-245 FROM N.A.
RC   STRAIN=NCTC 7273;
RA   Szabo E.A., Pemberton J.M., Desmarchelier P.M.;
```

```
RL   Submitted (APR-1992) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 633-993 FROM N.A.
RC   STRAIN=NCTC 7273;
RX   MEDLINE=94013372; PubMed=8408542;
RA   Campbell K., East A.K., Collins M.D.;
RT   "Gene probes for identification of the botulinal neurotoxin gene and
RT   specific identification of neurotoxin types B, E, and F.";
RL   J. Clin. Microbiol. 31:2255-2262(1993).
RN   [4]
RP   SEQUENCE OF 1-44 AND 441-466.
RC   STRAIN=657;
RX   MEDLINE=89000987; PubMed=3139097;
RA   Dasgupta B.R., Datta A.;
RT   "Botulinum neurotoxin type B (strain 657): partial sequence and
RT   similarity with tetanus toxin.";
RL   Biochimie 70:811-817(1988).
RN   [5]
RP   SEQUENCE OF 1-16 AND 441-458.
RC   STRAIN=OKRA;
RX   MEDLINE=85197963; PubMed=3888113;
RA   Schmidt J.J., Sathyamoorthy V., Dasgupta B.R.;
RT   "Partial amino acid sequences of botulinum neurotoxins types B and
RT   E.";
RL   Arch. Biochem. Biophys. 238:544-548(1985).
RN   [6]
RP   IDENTIFICATION AS ZINC-PROTEASE.
RX   MEDLINE=93054694; PubMed=1429690;
RA   Schiavo G., Rossetto O., Santucci A., Dasgupta B.R., Montecucco C.;
RT   "Botulinum neurotoxins are zinc proteins.";
RL   J. Biol. Chem. 267:23479-23483(1992).
RN   [7]
RP   IDENTIFICATION OF SUBSTRATE.
RX   MEDLINE=93063293; PubMed=1331807;
RA   Schiavo G., Benfenati F., Poulain B., Rossetto O., de Laureto P.P.,
RA   Dasgupta B.R., Montecucco C.;
RT   "Tetanus and botulinum-B neurotoxins block neurotransmitter release
RT   by proteolytic cleavage of synaptobrevin.";
RL   Nature 359:832-835(1992).
CC   -!- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC       RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC       AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC       WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC       INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC       ENDOPEPTIDASE THAT CLEAVES THE 76-GLN-|-PHE-77 BOND OF
CC       SYNAPTOBREVIN-2.
CC   -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC       neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
CC       detected action on small molecule substrates.
CC   -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC   -!- SUBUNIT: Disulfide-linked heterodimer of a light chain (L) and a
CC       heavy chain (H). The light chain has the pharmacological activity,
CC       while the N- and C-terminal of the heavy chain mediate channel
CC       formation and toxin binding, respectively.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- MISCELLANEOUS: There are seven antigenically distinct forms of
CC       botulinum neurotoxin: Types A, B, C1, D, E, F, and G.
CC   -!- SIMILARITY: Belongs to peptidase family M27.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M81186; AAA23211.1; -.
DR   EMBL; Z11934; CAA77991.1; -.
DR   EMBL; X70817; CAA50148.1; -.
DR   PIR; A48940; A48940.
DR   PDB; 1EPW; 01-NOV-00.
DR   PDB; 1F31; 01-NOV-00.
DR   PDB; 1F82; 16-AUG-00.
DR   PDB; 1F83; 16-AUG-00.
DR   PDB; 1FQH; 06-DEC-00.
DR   PDB; 1G9A; 13-NOV-02.
DR   PDB; 1G9B; 13-NOV-02.
DR   PDB; 1G9C; 13-NOV-02.
DR   PDB; 1G9D; 13-NOV-02.
DR   PDB; 1I1E; 21-NOV-01.
DR   MEROPS; M27.002; -.
DR   InterPro; IPR008985; ConA_like_lec_gl.
DR   InterPro; IPR002160; Kunitz_legume.
DR   InterPro; IPR006025; Pept_M_Zn_BS.
DR   InterPro; IPR000395; Peptidase_M27.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
KW   Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc;
KW   3D-structure.
FT   INIT_MET      0      0
FT   CHAIN         1    440       BOTULINUM NEUROTOXIN B, LIGHT-CHAIN.
FT   CHAIN       441   1290       BOTULINUM NEUROTOXIN B, HEAVY-CHAIN.
FT   METAL       229    229       ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT_SITE    230    230       BY SIMILARITY.
FT   METAL       233    233       ZINC (CATALYTIC) (BY SIMILARITY).
FT   DISULFID    436    445       INTERCHAIN (PROBABLE).
FT   CONFLICT     29     29       T -> M (IN REF. 4).
FT   CONFLICT    217    217       R -> G (IN REF. 2).
FT   CONFLICT    224    224       A -> S (IN REF. 2).
FT   CONFLICT    463    463       S -> R (IN REF. 4).
SQ   SEQUENCE   1290 AA;  150670 MW;  D21746E2C024DF43 CRC64;

  Query Match             31.5%;  Score 63;  DB 1;  Length 1290;
  Best Local Similarity   62.5%;  Pred. No. 0.98;
  Matches   10;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         15 SLFNNFTVSFWLRVPK 30
              |:| :|:||||:|:||
Db        920 SVFLDFSVSFWIRIPK 935


RESULT 6
VP2_AHSV6
ID   VP2_AHSV6      STANDARD;      PRT;  1051 AA.
AC   O71024;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Outer capsid protein VP2.
GN   S2 OR L2.
OS   African horse sickness virus 6 (AHSV-6) (African horse sickness virus
OS   (serotype 6)).
OC   Viruses; dsRNA viruses; Reoviridae; Orbivirus.
OX   NCBI_TaxID=86060;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98278331; PubMed=9617769;
RA   Williams C.F., Inoue T., Lucus A.-M., Zanotto P., Roy P.;
RT   "The complete sequence of four major structural proteins of African
RT   horse sickness virus serotype 6: evolutionary relationships within
RT   and between the orbiviruses.";
RL   Virus Res. 53:53-73(1998).
CC   -!- FUNCTION: THE VP2 PROTEIN IS ONE OF THE TWO PROTEINS (WITH VP5)
CC       WHICH CONSTITUTE THE VIRUS PARTICLE OUTER CAPSID. IT IS THE
CC       MAJOR TARGET OF THE HOST IMMUNOGENIC RESPONSE.
CC   -!- SIMILARITY: Belongs to the reoviruses VP2 protein family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
```

```
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AF021235; AAC40994.1; -.
DR   InterPro; IPR001742; Orbi_VP2.
DR   Pfam; PF00898; Orbi_VP2; 1.
KW   Coat protein.
SQ   SEQUENCE   1051 AA;  122326 MW;  2B04DB9E389F4B5F CRC64;

  Query Match             31.0%;  Score 62;  DB 1;  Length 1051;
  Best Local Similarity   40.7%;  Pred. No. 1.1;
  Matches   11;  Conservative    6;  Mismatches   10;  Indels    0;  Gaps    0;

Qy         11 SSGPSLFNNFTVSFWLRVPKVSASHLE 37
              : | : |:   ||:| || |::  |||
Db        630 TEGVTYFSKRFVSYWYRVEKITTKHLE 656


RESULT 7
BXA2_CLOBO
ID   BXA2_CLOBO     STANDARD;      PRT;  1295 AA.
AC   Q45894; P77780;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Botulinum neurotoxin type A precursor (EC 3.4.24.69) (BoNT/A)
DE   (Bontoxilysin A) (BOTOX) [Contains: Botulinum neurotoxin A, light-
DE   chain; Botulinum neurotoxin A, heavy-chain].
GN   BOTA OR BNA OR ATX.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Kyoto-F;
RX   MEDLINE=94143603; PubMed=8310180;
RA   Willems A., East A.K., Lawson P.A., Collins M.D.;
RT   "Sequence of the gene coding for the neurotoxin of Clostridium
RT   botulinum type A associated with infant botulism: comparison with
RT   other clostridial neurotoxins.";
RL   Res. Microbiol. 144:547-556(1993).
RN   [2]
RP   SEQUENCE OF 1-65 FROM N.A.
RC   STRAIN=Kyoto-F;
RX   MEDLINE=97016817; PubMed=8863443;
RA   East A.K., Bhandari M., Stacey J.M., Campbell K.D., Collins M.D.;
RT   "Organization and phylogenetic interrelationships of genes encoding
RT   components of the botulinum toxin complex in proteolytic Clostridium
RT   botulinum types A, B, and F: evidence of chimeric sequences in the
RT   gene encoding the nontoxic nonhemagglutinin component.";
RL   Int. J. Syst. Bacteriol. 46:1105-1112(1996).
CC   -!- FUNCTION: Inhibits acetylcholine release. The botulinum toxin
CC       binds with high affinity to peripheral neuronal presynaptic
CC       membrane, is then internalized by receptor-mediated endocytosis.
CC       The C-terminus of the heavy chain (H) is responsible for the
CC       adherence of the toxin to the cell surface while the N-terminus
CC       mediates transport of the light chain from the endocytic vesicle
CC       to the cytosol. After translocation, the light chain (L)
CC       hydrolyzes the 197-Gln-|-Arg-198 bond in SNAP-25, thereby blocking
CC       neurotransmitter release. Inhibition of acetylcholine release
CC       results in flaccid paralysis, with frequent heart or respiratory
CC       failure (By similarity).
CC   -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC       neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
CC       detected action on small molecule substrates.
CC   -!- SUBUNIT: Disulfide-linked heterodimer of a light chain (L) and a
CC       heavy chain (H) (By similarity).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- MISCELLANEOUS: There are seven antigenically distinct forms of
CC       botulinum neurotoxin: Types A, B, C1, D, E, F, and G.
CC   -!- SIMILARITY: Belongs to peptidase family M27.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; X73423; CAA51824.1; -.
DR   EMBL; X87974; CAA61234.1; -.
DR   PIR; I40645; I40645.
DR   HSSP; P10845; 3BTA.
DR   MEROPS; M27.002; -.
DR   InterPro; IPR008985; ConA_like_lec_gl.
DR   InterPro; IPR002160; Kunitz_legume.
DR   InterPro; IPR006025; Pept_M_Zn_BS.
DR   InterPro; IPR000395; Peptidase_M27.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; FALSE_NEG.
KW   Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
FT   INIT_MET      0      0       BY SIMILARITY.
FT   CHAIN         1    447       BOTULINUM NEUROTOXIN A, LIGHT-CHAIN.
FT   CHAIN       448   1295       BOTULINUM NEUROTOXIN A, HEAVY-CHAIN.
FT   METAL       222    222       ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT_SITE    223    223       BY SIMILARITY.
FT   METAL       226    226       ZINC (CATALYTIC) (BY SIMILARITY).
FT   DISULFID    429    453       INTERCHAIN (BY SIMILARITY).
FT   DISULFID   1234   1279       BY SIMILARITY.
FT   TRANSMEM    626    646       POTENTIAL.
FT   TRANSMEM    655    675       POTENTIAL.
SQ   SEQUENCE   1295 AA;  149279 MW;  5DA04A13D98D6372 CRC64;

  Query Match             31.0%;  Score 62;  DB 1;  Length 1295;
  Best Local Similarity   50.0%;  Pred. No. 1.4;
  Matches    8;  Conservative    6;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         15 SLFNNPTVSFWLRVPK 30
              |:: ||: |||:::||
Db        935 SMYENFSTSFWIKIPK 950


RESULT 8
GON1_TUPGB
ID   GON1_TUPGB     STANDARD;      PRT;    92 AA.
AC   Q95335;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Progonadoliberin I precursor [Contains: Gonadoliberin I (LH-RH I)
DE   (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
DE   hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I].
GN   GNRH1 OR GNRH.
OS   Tupaia glis belangeri (Common tree shrew).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Scandentia; Tupaiidae; Tupaia.
OX   NCBI_TaxID=37347;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Hypothalamus;
RX   MEDLINE=97079639; PubMed=8921350;
RA   Kasten T.L., White S.A., Norton T.T., Bond C.T., Adelman J.P.,
RA   Fernald R.D.;
RT   "Characterization of two new preproGnRH mRNAs in the tree shrew:
RT   first direct evidence for mesencephalic GnRH gene expression in a
RT   placental mammal.";
RL   Gen. Comp. Endocrinol. 104:7-19(1996).
CC   -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
CC       the secretion of both luteinizing and follicle-stimulating
CC       hormones.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the GnRH family.
CC   ---------------------------------------------------------------------------
```

```
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; U63326; AAB16837.1; -.
DR   InterPro; IPR002012; GnRH.
DR   InterPro; IPR004079; GonadoliberinI.
DR   Pfam; PF00446; GnRH; 1.
DR   PRINTS; PR01541; GONADOLIBRNI.
DR   PROSITE; PS00473; GNRH; 1.
KW   Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW   Placenta; Signal; Pyrrolidone carboxylic acid.
FT   SIGNAL        1     23       BY SIMILARITY.
FT   CHAIN        24     92       PROGONADOLIBERIN I.
FT   PEPTIDE      24     33       GONADOLIBERIN I.
FT   PEPTIDE      37     92       GNRH-ASSOCIATED PEPTIDE I.
FT   ACT_SITE     26     26       APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT                                ACTIVITY.
FT   MOD_RES      24     24       PYRROLIDONE CARBOXYLIC ACID (BY
FT                                SIMILARITY).
FT   MOD_RES      33     33       AMIDATION (G-34 PROVIDE AMIDE GROUP) (BY
FT                                SIMILARITY).
SQ   SEQUENCE   92 AA;  10197 MW;  4FDBF2C58CF5F63B CRC64;

  Query Match             30.0%;  Score 60;  DB 1;  Length 92;
  Best Local Similarity   52.4%;  Pred. No. 0.12;
  Matches   11;  Conservative    4;  Mismatches    4;  Indels    2;  Gaps    1;

Qy          2 HWSYGLRPGS--SGPSLFNNF 20
              |||||||||   :  :| ::|
Db         25 HWSYGLRPGGKRNAENLIDSF 45


RESULT 9
GON1_MACMU
ID   GON1_MACMU     STANDARD;      PRT;    67 AA.
AC   P55247;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Progonadoliberin I precursor [Contains: Gonadoliberin I (LH-RH I)
DE   (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
DE   hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I]
DE   (Fragment).
GN   GNRH1 OR GNRH OR LHRH.
OS   Macaca mulatta (Rhesus macaque).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9544;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Hypothalamus;
RX   MEDLINE=95124501; PubMed=7545971;
RA   Ma Y.J., Costa M.E., Ojeda S.R.;
RT   "Developmental expression of the genes encoding transforming growth
RT   factor alpha and its receptor in the hypothalamus of female rhesus
RT   macaques.";
RL   Neuroendocrinology 60:346-359(1994).
CC   -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
CC       the secretion of both luteinizing and follicle-stimulating
CC       hormones.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the GnRH family.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; S75918; AAB33096.1; -.
DR   PIR; I78541; I78541.
DR   InterPro; IPR002012; GnRH.
DR   InterPro; IPR004079; GonadoliberinI.
DR   Pfam; PF00446; GnRH; 1.
DR   PRINTS; PR01541; GONADOLIBRNI.
DR   PROSITE; PS00473; GNRH; 1.
KW   Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW   Signal; Pyrrolidone carboxylic acid.
FT   NON_TER       1      1
FT   SIGNAL       <1      5       BY SIMILARITY.
FT   CHAIN         6     >67      PROGONADOLIBERIN I.
FT   PEPTIDE       6     15       GONADOLIBERIN I.
FT   PEPTIDE      19     >67      GNRH-ASSOCIATED PEPTIDE I.
FT   ACT_SITE      8      8       APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT                                ACTIVITY (BY SIMILARITY).
FT   MOD_RES       6      6       PYRROLIDONE CARBOXYLIC ACID (BY
FT                                SIMILARITY).
FT   MOD_RES      15     15       AMIDATION (G-16 PROVIDE AMIDE GROUP) (BY
FT                                SIMILARITY).
FT   NON_TER      67     67
SQ   SEQUENCE   67 AA;  7573 MW;  505394DAA261A3F2 CRC64;

  Query Match             29.5%;  Score 59;  DB 1;  Length 67;
  Best Local Similarity   52.4%;  Pred. No. 0.11;
  Matches   11;  Conservative    3;  Mismatches    5;  Indels    2;  Gaps    1;

Qy          2 HWSYGLRPGS--SGPSLFNNF 20
              |||||||||   :| ::|
Db          7 HWSYGLRPGGKRDAENLMDSF 27


RESULT 10
GON1_HUMAN
ID   GON1_HUMAN     STANDARD;      PRT;    92 AA.
AC   P01148;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-APR-1988 (Rel. 07, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Progonadoliberin I precursor [Contains: Gonadoliberin I (LH-RH I)
DE   (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
DE   hormone I) (GnRH I) (Luliberin I) (Gonadorelin); GnRH-associated
DE   peptide I].
GN   GNRH1 OR GNRH OR LHRH.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89366682; PubMed=2671939;
RA   Hayflick J.S., Adelman J.P., Seeburg P.H.;
RT   "The complete nucleotide sequence of the human gonadotropin-releasing
RT   hormone gene.";
RL   Nucleic Acids Res. 17:6403-6403(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86094338; PubMed=2867548;
RA   Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;
RT   "Isolation of the gene and hypothalamic cDNA for the common precursor
RT   of gonadotropin-releasing hormone and prolactin release-inhibiting
RT   factor in human and rat.";
RL   Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).
RN   [3]
RP   SEQUENCE FROM N.A., AND VARIANT SER-16.
RX   MEDLINE=85012739; PubMed=6090951;
RA   Seeburg P.H., Adelman J.P.;
RT   "Characterization of cDNA for precursor of human luteinizing hormone
RT   releasing hormone.";
```

```
RL   Nature 311:666-668(1984).
RN   [4]
RP   SEQUENCE OF 24-33.
RX   MEDLINE=83126573; PubMed=6760865;
RA   Tan L., Rousseau P.;
RT   "The chemical identity of the immunoreactive LHRH-like peptide
RT   biosynthesized in the human placenta.";
RL   Biochem. Biophys. Res. Commun. 109:1061-1071(1982).
RN   [5]
RP   VARIANT SER-16.
RX   MEDLINE=99318093; PubMed=10391209;
RA   Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
RA   Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,
RA   Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,
RA   Lander E.S.;
RT   "Characterization of single-nucleotide polymorphisms in coding regions
RT   of human genes.";
RL   Nat. Genet. 22:231-238(1999).
RN   [6]
RP   ERRATUM.
RA   Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
RA   Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,
RA   Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,
RA   Lander E.S.;
RL   Nat. Genet. 23:373-373(1999).
CC   -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
CC       the secretion of both luteinizing and follicle-stimulating
CC       hormones.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- PHARMACEUTICAL: Available under the names Factrel (Ayerst Labs),
CC       Lutrepulse or Lutrelef (Ferring Pharmaceuticals) and Relisorm
CC       (Serono).
CC   -!- SIMILARITY: Belongs to the GnRH family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X01059; CAA25526.1; -.
DR   EMBL; M12578; AAA35916.1; -.
DR   EMBL; X15215; CAA33285.1; -.
DR   PIR; S05308; RHHUG.
DR   Genew; HGNC:4419; GNRH1.
DR   MIM; 152760; -.
DR   GO; GO:0005625; C:soluble fraction; TAS.
DR   GO; GO:0005183; F:luteinizing hormone-releasing factor activity; TAS.
DR   GO; GO:0007267; P:cell-cell signaling; TAS.
DR   GO; GO:0007275; P:development; TAS.
DR   GO; GO:0008285; P:negative regulation of cell proliferation; TAS.
DR   GO; GO:0007165; P:signal transduction; TAS.
DR   InterPro; IPR002012; GnRH.
DR   InterPro; IPR004079; GonadoliberinI.
DR   Pfam; PF00446; GnRH; 1.
DR   PRINTS; PR01541; GONADOLIBRNI.
DR   PROSITE; PS00473; GNRH; 1.
KW   Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW   Placenta; Pharmaceutical; Signal; Polymorphism;
KW   Pyrrolidone carboxylic acid.
FT   SIGNAL        1     23
FT   CHAIN        24     92       PROGONADOLIBERIN I.
FT   PEPTIDE      24     33       GONADOLIBERIN I.
FT   PEPTIDE      37     92       GNRH-ASSOCIATED PEPTIDE I.
FT   ACT_SITE     26     26       APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT                                ACTIVITY.
FT   MOD_RES      24     24       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES      33     33       AMIDATION (G-34 PROVIDE AMIDE GROUP).
FT   VARIANT      16     16       W -> S (in dbSNP:6185).
FT                                /FTId=VAR_013943.
SQ   SEQUENCE   92 AA;  10380 MW;  30A72221B076FA79 CRC64;
```

```
  Query Match           29.5%;  Score 59;  DB 1;  Length 92;
  Best Local Similarity  52.4%;  Pred. No. 0.16;
  Matches   11;  Conservative    3;  Mismatches    5;  Indels    2;  Gaps    1;

Qy          2 HWSYGLRPGS--SGPSLFNNF 20
              |||||||||      :| ::|
Db         25 HWSYGLRPGGKRDAENLIDSF 45


RESULT 11
GON1_SHEEP
ID   GON1_SHEEP     STANDARD;      PRT;    61 AA.
AC   Q28588;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Progonadoliberin I precursor [Contains: Gonadoliberin I (LH-RH I)
DE   (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
DE   hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I]
DE   (Fragment).
GN   GNRH1 OR GNRH OR LHRH.
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Caprinae; Ovis.
OX   NCBI_TaxID=9940;
RN   [1]
RP   SEQUENCE OF 12-61 FROM N.A.
RC   STRAIN=Western range; TISSUE=Hypothalamus;
RA   Rodriguez R.E., Wise M.E.;
RL   Submitted (OCT-1993) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 1-10.
RX   MEDLINE=72094314; PubMed=4550508;
RA   Burgus R., Butcher M., Amoss M., Ling N., Monahan M., Rivier J.,
RA   Fellows R., Blackwell R., Vale W., Guillemin R.;
RT   "Primary structure of the ovine hypothalamic luteinizing hormone-
RT   releasing factor (LRF) (LH-hypothalamus-LRF-gas chromatography-mass
RT   spectrometry-decapeptide-Edman degradation).";
RL   Proc. Natl. Acad. Sci. U.S.A. 69:278-282(1972).
CC   -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
CC       the secretion of both luteinizing and follicle-stimulating
CC       hormones.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the GnRH family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U02517; AAA03433.1; -.
DR   InterPro; IPR002012; GnRH.
DR   InterPro; IPR004079; GonadoliberinI.
DR   Pfam; PF00446; GnRH; 1.
DR   PRINTS; PR01541; GONADOLIBRNI.
DR   PROSITE; PS00473; GNRH; 1.
KW   Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW   Placenta; Pyrrolidone carboxylic acid.
FT   NON_TER       1      1
FT   CHAIN         1    >61       PROGONADOLIBERIN I.
FT   PEPTIDE       1     10       GONADOLIBERIN I.
FT   PEPTIDE      14    >61       GNRH-ASSOCIATED PEPTIDE I.
FT   ACT_SITE      3      3       APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT                                ACTIVITY.
FT   MOD_RES       1      1       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES      10     10       AMIDATION (G-11 PROVIDE AMIDE GROUP).
FT   NON_TER      61     61
SQ   SEQUENCE   61 AA;  6828 MW;  63962A1AE319B8F0 CRC64;
```

```
  Query Match           29.0%;  Score 58;  DB 1;  Length 61;
  Best Local Similarity  100.0%;  Pred. No. 0.14;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 HWSYGLRPG 10
              |||||||||
Db          2 HWSYGLRPG 10


RESULT 12
GON1_MESAU
ID   GON1_MESAU     STANDARD;      PRT;    63 AA.
AC   O09163;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Progonadoliberin I precursor [Contains: Gonadoliberin I (LH-RH I)
DE   (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
DE   hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I]
DE   (Fragment).
GN   GNRH1 OR GNRH OR LHRH.
OS   Mesocricetus auratus (Golden hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC   Mesocricetus.
OX   NCBI_TaxID=10036;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Jansen H.T., Stevens P.J., Zeitler P., Lehman M.N.;
RL   Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
CC       the secretion of both luteinizing and follicle-stimulating
CC       hormones.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the GnRH family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U91938; AAB51302.1; -.
DR   InterPro; IPR002012; GnRH.
DR   InterPro; IPR004079; GonadoliberinI.
DR   Pfam; PF00446; GnRH; 1.
DR   PRINTS; PR01541; GONADOLIBRNI.
DR   PROSITE; PS00473; GNRH; 1.
KW   Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW   Placenta; Pyrrolidone carboxylic acid.
FT   NON_TER       1      1
FT   CHAIN         1    >63       PROGONADOLIBERIN I.
FT   PEPTIDE       1     10       GONADOLIBERIN I.
FT   PEPTIDE      14    >63       GNRH-ASSOCIATED PEPTIDE I (BY
FT                                 SIMILARITY).
FT   ACT_SITE      3      3       APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT                                 ACTIVITY (BY SIMILARITY).
FT   MOD_RES       1      1       PYRROLIDONE CARBOXYLIC ACID (BY
FT                                 SIMILARITY).
FT   MOD_RES      10     10       AMIDATION (G-11 PROVIDE AMIDE GROUP) (BY
FT                                 SIMILARITY).
FT   NON_TER      63     63
SQ   SEQUENCE   63 AA;  7370 MW;  FC94995676F77180 CRC64;

  Query Match           29.0%;  Score 58;  DB 1;  Length 63;
  Best Local Similarity  100.0%;  Pred. No. 0.15;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 HWSYGLRPG 10
              |||||||||
Db          2 HWSYGLRPG 10


RESULT 13
GON1_XENLA
ID   GON1_XENLA     STANDARD;      PRT;    89 AA.
AC   P45656;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Gonadoliberin I precursor (Gonadotropin-releasing hormone I) (GnRH-I)
DE   (LH-RH) (Luliberin I).
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Forebrain;
RX   MEDLINE=94185563; PubMed=8137750;
RA   Hayes W.P., Wray S., Battey J.F.;
RT   "The frog gonadotropin-releasing hormone-I (GnRH-I) gene has a
RT   mammalian-like expression pattern and conserved domains in
RT   GnRH-associated peptide, but brain onset is delayed until
RT   metamorphosis.";
RL   Endocrinology 134:1835-1844(1994).
CC   -!- FUNCTION: Stimulates the secretion of gonadotropins.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the GnRH family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L28040; AAA49728.1; -.
DR   PIR; I51423; I51423.
DR   InterPro; IPR002012; GnRH.
DR   InterPro; IPR004079; GonadoliberinI.
DR   Pfam; PF00446; GnRH; 1.
DR   PRINTS; PR01541; GONADOLIBRNI.
DR   PROSITE; PS00473; GNRH; 1.
KW   Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW   Signal; Pyrrolidone carboxylic acid.
FT   SIGNAL        1     23
FT   CHAIN        24     89       PROGONADOLIBERIN I.
FT   PEPTIDE      24     33       GONADOLIBERIN I.
FT   PEPTIDE      37     89       GONADOTROPIN-RELEASING HORMONE ASSOCIATED
FT                                 PEPTIDE.
FT   PEPTIDE      37     85       GNRH-ASSOCIATED PEPTIDE I (GAP).
FT   MOD_RES      24     24       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES      33     33       AMIDATION (G-34 PROVIDE AMIDE GROUP).
SQ   SEQUENCE   89 AA;  10246 MW;  6F4F36FBAE0D4284 CRC64;

  Query Match           29.0%;  Score 58;  DB 1;  Length 89;
  Best Local Similarity  100.0%;  Pred. No. 0.22;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 HWSYGLRPG 10
              |||||||||
Db         25 HWSYGLRPG 33


RESULT 14
GON1_MOUSE
ID   GON1_MOUSE     STANDARD;      PRT;    90 AA.
AC   P13562;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
```

```
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Progonadoliberin I precursor [Contains: Gonadoliberin I (LH-RH I)
DE   (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
DE   hormone I) (GnRH I) (Luliberin I); Prolactin release-inhibiting factor
DE   I].
GN   GNRH1 OR GNRH.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87069928; PubMed=3024317;
RA   Mason A.J., Hayflick J.S., Zoeller R.T., Young W.S. III,
RA   Phillips H.S., Nikolics K., Seeburg P.H.;
RT   "A deletion truncating the gonadotropin-releasing hormone gene is
RT   responsible for hypogonadism in the hpg mouse.";
RL   Science 234:1366-1371(1986).
CC   -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
CC       the secretion of both luteinizing and follicle-stimulating
CC       hormones.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the GnRH family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M14872; AAA37717.1; -.
DR   PIR; A47578; RHMSG.
DR   MGD; MGI:95789; Gnrh.
DR   InterPro; IPR002012; GnRH.
DR   InterPro; IPR004079; GonadoliberinI.
DR   Pfam; PF00446; GnRH; 1.
DR   PRINTS; PR01541; GONADOLIBRNI.
DR   PROSITE; PS00473; GNRH; 1.
KW   Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW   Placenta; Signal; Pyrrolidone carboxylic acid.
FT   SIGNAL        1     21
FT   CHAIN        22     90       PROGONADOLIBERIN I.
FT   PEPTIDE      22     31       GONADOLIBERIN I.
FT   PEPTIDE      35     90       PROLACTIN RELEASE-INHIBITING FACTOR I.
FT   ACT_SITE     24     24       APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT                                ACTIVITY.
FT   MOD_RES      22     22       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES      31     31       AMIDATION (G-32 PROVIDE AMIDE GROUP).
SQ   SEQUENCE   90 AA;  10337 MW;  1C0766FA4826B4D9 CRC64;

  Query Match             29.0%;  Score 58;  DB 1;  Length 90;
  Best Local Similarity   100.0%;  Pred. No. 0.22;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 HWSYGLRPG 10
              |||||||||
Db         23 HWSYGLRPG 31


RESULT 15
GON1_RANCA
ID   GON1_RANCA     STANDARD;      PRT;    90 AA.
AC   Q90Y63;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Progonadoliberin I precursor [Contains: Gonadoliberin I (LHRH I)
DE   (Luteinizing hormone releasing hormone I) (Gonadotropin releasing
DE   hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I (GAP1)].
GN   GNRH1 OR GNRH.
OS   Rana catesbeiana (Bull frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Ranidae; Rana.
OX   NCBI_TaxID=8400;
RN   [1]
RP   SEQUENCE FROM N.A., TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.
RC   TISSUE=Forebrain;
RX   MEDLINE=21102951; PubMed=11170016;
RA   Wang L., Yoo M.S., Kang H.M., Im W.B., Choi H.S., Bogerd J.,
RA   Kwon H.B.;
RT   "Cloning and characterization of cDNAs encoding the GnRH1 and GnRH2
RT   precursors from bullfrog (Rana catesbeiana).";
RL   J. Exp. Zool. 289:190-201(2001).
CC   -!- FUNCTION: Stimulates the secretion of gonadotropins (By
CC       similarity).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Forebrain.
CC   -!- DEVELOPMENTAL STAGE: Expressed at significantly higher levels
CC       during post-breeding. Not expressed in pituitary.
CC   -!- SIMILARITY: Belongs to the GnRH family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF188754; AAL05972.1; -.
DR   GO; GO:0005576; C:extracellular; NAS.
DR   GO; GO:0005183; F:luteinizing hormone-releasing factor activity; NAS.
DR   GO; GO:0009755; P:hormone mediated signaling; NAS.
DR   GO; GO:0000003; P:reproduction; NAS.
DR   InterPro; IPR002012; GnRH.
DR   InterPro; IPR004079; GonadoliberinI.
DR   Pfam; PF00446; GnRH; 1.
DR   PRINTS; PR01541; GONADOLIBRNI.
DR   PROSITE; PS00473; GNRH; 1.
KW   Cleavage on pair of basic residues; Hormone; Amidation; Signal;
KW   Pyrrolidone carboxylic acid.
FT   SIGNAL        1     24       POTENTIAL.
FT   CHAIN        25     90       PROGONADOLIBERIN I.
FT   PEPTIDE      25     34       GONADOLIBERIN I.
FT   PEPTIDE      38     86       GNRH-ASSOCIATED PEPTIDE I (BY
FT                                SIMILARITY).
FT   MOD_RES      25     25       PYRROLIDONE CARBOXYLIC ACID (BY
FT                                SIMILARITY).
FT   MOD_RES      34     34       AMIDATION (G-35 PROVIDE AMIDE GROUP) (BY
FT                                SIMILARITY).
SQ   SEQUENCE   90 AA;  10291 MW;  317203B4E3DA2FE7 CRC64;

  Query Match             29.0%;  Score 58;  DB 1;  Length 90;
  Best Local Similarity   100.0%;  Pred. No. 0.22;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 HWSYGLRPG 10
              |||||||||
Db         26 HWSYGLRPG 34


Search completed: March 10, 2004, 09:13:56
Job time : 7.76654 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:54 ; Search time 11.5175 Seconds
(without alignments)
309.015 Million cell updates/sec

Title: US-09-848-834A-14
Perfect score: 200
Sequence: 1 XHWSYGLRPGSSGPSLFNNFTVSFWLRVPKVSASHLE 37

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	114	57.0	1315	1	BTCLTN	tentoxilysin (EC 3
2	67	33.5	1297	2	S39791	neurotoxin - Clost
3	65	32.5	1296	1	BTCLAB	bontoxilysin (EC 3
4	64	32.0	1268	2	S33411	botulinum neurotox
5	64	32.0	1291	2	I40631	non-proteolytic bo
6	63	31.5	369	2	S48109	neurotoxin type F
7	63	31.5	1274	2	I40813	neurotoxin type F
8	63	31.5	1291	1	A48940	bontoxilysin (EC 3
9	62	31.0	1296	2	I40645	botulinum neurotox
10	61	30.5	366	2	S48110	neurotoxin type F
11	59	29.5	67	2	I78541	gonadoliberin prec
12	59	29.5	92	1	RHHUG	gonadoliberin prec
13	58	29.0	10	1	RHPGG	gonadoliberin - pi
14	58	29.0	10	1	RHSHG	gonadoliberin - sh
15	58	29.0	89	2	I51423	gonadoliberin prec
16	58	29.0	90	1	RHMSG	gonadoliberin prec
17	58	29.0	92	1	RHRTG	gonadoliberin prec
18	57.5	28.7	352	1	VVVP24	coat protein VP2 -
19	56.5	28.2	98	2	I50739	gonadotropin-relea
20	56	28.0	92	2	I50644	gonadoliberin I pr
21	56	28.0	367	2	S48106	neurotoxin type E
22	56	28.0	1251	2	JH0256	botulinum neurotox
23	56	28.0	1252	2	S21178	botulinum neurotox
24	56	28.0	1291	2	S46431	botulinum neurotox
25	56	28.0	1291	2	A49777	botulinum neurotox
26	54.5	27.3	91	2	JC7393	medaka-type gonado
27	54.5	27.3	464	1	MNVUWC	nonstructural prot
28	54	27.0	10	1	RHAQ1	gonadoliberin I -
29	53.5	26.8	251	2	AD1669	3'-exo-deoxyribonu
30	53	26.5	449	2	S23158	nucleocapsid prote
31	53	26.5	1196	2	JQ1467	toxin, nontoxic co
32	53	26.5	1196	2	S46430	botulinum neurotox
33	53	26.5	1285	2	S70582	botulinum neurotox
34	52.5	26.2	251	2	AF1297	3'-exo-deoxyribonu
35	52.5	26.2	1999	2	AB2018	hypothetical prote
36	52	26.0	210	2	T18703	hypothetical prote
37	52	26.0	467	1	MNVUW1	nonstructural prot
38	52	26.0	916	2	T04752	aspartate kinase (
39	51.5	25.8	564	2	I48776	spermatogenesis re
40	51.5	25.8	1033	2	I48775	Smcx protein (esca
41	51.5	25.8	1560	2	I54361	SMCX protein - hum
42	51	25.5	436	2	D84782	probable proline t
43	51	25.5	637	2	F90257	hypothetical prote
44	51	25.5	1261	2	AI3471	5-methyltetrahydro
45	50.5	25.2	230	2	T27498	hypothetical prote

ALIGNMENTS

RESULT 1
BTCLTN
tentoxilysin (EC 3.4.24.68) precursor - Clostridium tetani
N;Alternate names: tetanus neurotoxin
C;Species: Clostridium tetani
C;Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 03-Jun-2002
C;Accession: A25689; A25757; A25194; B25194; A60759; S69348; S09364
R;Eisel, U.; Jarausch, W.; Goretzki, K.; Henschen, A.; Engels, J.; Weller, U.; Hudel,
EMBO J. 5, 2495-2502, 1986
A;Title: Tetanus toxin: primary structure, expression in E. coli, and homology with bo
A;Reference number: A25689; MUID:87053814; PMID:3536478
A;Accession: A25689
A;Molecule type: DNA
A;Residues: 1-1315 <EIS>
A;Cross-references: GB:X04436; NID:g40769; PIDN:CAA28033.1; PID:g40770
R;Fairweather, N.F.; Lyness, V.A.
Nucleic Acids Res. 14, 7809-7812, 1986
A;Title: The complete nucleotide sequence of tetanus toxin.
A;Reference number: A25757; MUID:87040747; PMID:3774547
A;Accession: A25757
A;Molecule type: DNA
A;Residues: 1-1315 <FAI>
A;Cross-references: GB:X06214; NID:g40773; PIDN:CAA29564.1; PID:g40774
A;Experimental source: strain CN3911
R;Fairweather, N.F.; Lyness, V.A.; Pickard, D.J.; Allen, G.; Thomson, R.O.
J. Bacteriol. 165, 21-27, 1986
A;Title: Cloning, nucleotide sequencing, and expression of tetanus toxin fragment C in
A;Reference number: A25194; MUID:86085672; PMID:3510187
A;Accession: A25194
A;Molecule type: DNA
A;Residues: 743-1315 <FA2>
A;Cross-references: GB:M12739; NID:g144920; PIDN:AAA23282.1; PID:g144921
A;Accession: B25194
A;Molecule type: protein
A;Residues: 865-894 <FA3>
R;Matsuda, M.; Lei, D.L.; Sugimoto, N.; Ozutsumi, K.; Okabe, T.
Infect. Immun. 57, 3588-3593, 1989
A;Title: Isolation, purification, and characterization of fragment B, the NH-2-termina
A;Reference number: A60759; MUID:90035436; PMID:2478476
A;Accession: A60759
A;Molecule type: protein
A;Residues: 461-475 <MAT>
R;Demotz, S.; Lanzavecchia, L.; Eisel, U.; Niemann, H.; Widmann, C.; Corradin, G.
J. Immunol. 142, 394-402, 1989
A;Title: Delineation of several DR-restricted tetanus toxin T cell epitopes.
A;Reference number: JS0098; MUID:89093918; PMID:2463305
A;Contents: annotation; epitope region
R;Schiavo, G.; Benfenati, F.; Poulain, B.; Rossetto, O.; de Laureto, P.P.; DasGupta, B
Nature 359, 832-835, 1992
A;Title: Tetanus and botulinum-B neurotoxins block neurotransmitter release by proteol
A;Reference number: S27125; MUID:93063293; PMID:1331807
A;Contents: annotation

R;de Filippis, V.; Vangelista, L.; Schiavo, G.; Tonello, F.; Montecucco, C.
Eur. J. Biochem. 229, 61-69, 1995
A;Title: Structural studies on the zinc-endopeptidase light chain of tetanus neurotoxin.
A;Reference number: S69348; MUID:95262688; PMID:7744050
A;Accession: S69348
A;Molecule type: protein
A;Residues: 2-31 <DEF>
C;Comment: The source of this protein was an extrachromosomal plasmid.
C;Comment: The precursor is cleaved by endogenous proteinase activity to form light (fra
dual chains are not toxic when separated). The amino end of the heavy chain (fragment B)
C;Comment: Fragment B forms ion channels in a lipid bilayer. Fragment C binds to ganglic
C;Comment: This potent neurotoxin binds to peripheral neuronal synapses, is internalized
presynaptic neurons. It inhibits neurotransmitter release by proteolytic cleavage of sy
C;Function:
A;Description: blocks neuroexocytosis via hydrolysis of a Gln-Phe peptide bond in synapt
C;Superfamily: tetanus toxin
C;Keywords: hydrolase; metalloproteinase; neurotoxin; transmembrane protein; zinc
F;2-457/Product: tentoxylysin light chain (fragment A) #status predicted <TTL>
F;461-1315/Product: tentoxylysin heavy chain (fragment B.C) #status experimental <TTH>
F;461-864/Domain: channel forming (fragment B) #status predicted <TXB>
F;865-1315/Domain: ganglioside binding (fragment C) #status predicted <TXC>
F;233,237/Binding site: zinc (His) #status predicted
F;234/Active site: Glu #status predicted

Query Match 57.0%; Score 114; DB 1; Length 1315;
Best Local Similarity 95.5%; Pred. No. 1.3e-07;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LFNNFTVSFWLRVPKVSASHLE 37
:|||||||||||||||||||||
Db 946 MFNNFTVSFWLRVPKVSASHLE 967

RESULT 2
S39791
neurotoxin - Clostridium botulinum
C;Species: Clostridium botulinum
C;Date: 07-Oct-1994 #sequence_revision 01-Dec-1995 #text_change 16-Jul-1999
C;Accession: S39791
R;Campbell, K.; Collins, M.D.; East, A.K.
Biochim. Biophys. Acta 1216, 487-491, 1993
A;Title: Nucleotide sequence of the gene coding for Clostridium botulinum (Clostridium a
A;Reference number: S39791; MUID:94092745; PMID:8268233
A;Accession: S39791
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1297 <CAM>
A;Cross-references: EMBL:X74162; NID:g441275; PIDN:CAA52275.1; PID:g441276
C;Superfamily: tetanus toxin
C;Keywords: neurotoxin

Query Match 33.5%; Score 67; DB 2; Length 1297;
Best Local Similarity 39.1%; Pred. No. 0.54;
Matches 9; Conservative 10; Mismatches 4; Indels 0; Gaps 0;

Qy 15 SLFNNFTVSFWLRVPKVSASHLE 37
|:|:||:::||:| || : : ::
Db 928 SMFDNFSINFWVRTPKYNNNDIQ 950

RESULT 3
BTCLAB
bontoxilysin (EC 3.4.24.69) A precursor - Clostridium botulinum
N;Alternate names: botulinum neurotoxin type A
C;Species: Clostridium botulinum
C;Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 18-Jun-1999
C;Accession: A35294; S09492; S68220; A33401; A53884; A60025; A27000
R;Binz, T.; Kurazono, H.; Wille, M.; Frevert, J.; Wernars, K.; Niemann, H.
J. Biol. Chem. 265, 9153-9158, 1990
A;Title: The complete sequence of botulinum neurotoxin type A and comparison with other
A;Reference number: A35294; MUID:90264400; PMID:2160960
A;Accession: A35294
A;Molecule type: DNA
A;Residues: 1-1296 <BIN>
A;Cross-references: GB:M30196; NID:g144864; PIDN:AAA23262.1; PID:g144865
A;Experimental source: strain 62A, subtype A
R;Thompson, D.E.; Brehm, J.K.; Oultram, J.D.; Swinfield, T.J.; Shone, C.C.; Atkinson, T
Eur. J. Biochem. 189, 73-81, 1990
A;Title: The complete amino acid sequence of the Clostridium botulinum type A neurotoxi
A;Reference number: S09492; MUID:90235864; PMID:2185020
A;Accession: S09492
A;Molecule type: DNA
A;Residues: 1,'Q',3-26,'V',28-1296 <THO>
A;Cross-references: EMBL:X52066; NID:g40381; PIDN:CAA36289.1; PID:g40382
A;Experimental source: NCTC 2916
R;Fujita, R.; Fujinaga, Y.; Inoue, K.; Nakajima, H.; Kumon, H.; Oguma, K.
FEBS Lett. 376, 41-44, 1995
A;Title: Molecular characterization of two forms of nontoxic-nonhemagglutinin component
A;Reference number: S67988; MUID:96096783; PMID:8521962
A;Accession: S68220
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-12 <FUJ>
A;Cross-references: EMBL:D67030; DDBJ:D50421; NID:g2160224
R;Betley, M.J.; Somers, E.; DasGupta, B.R.
Biochem. Biophys. Res. Commun. 162, 1388-1395, 1989
A;Title: Characterization of botulinum type A neurotoxin gene: delineation of the N-ter
A;Reference number: A33401; MUID:89350959; PMID:2669749
A;Accession: A33401
A;Molecule type: DNA
A;Residues: 1-35 <BET>
A;Cross-references: GB:M27892; NID:g144880; PIDN:AAA23269.1; PID:g551776
R;Gimenez, J.A.; DasGupta, B.R.
J. Protein Chem. 12, 351-363, 1993
A;Title: Botulinum type A neurotoxin digested with pepsin yields 132, 97, 72, 45, 42, a
A;Reference number: A53884; MUID:94000342; PMID:8397793
A;Accession: A53884
A;Status: preliminary
A;Molecule type: protein
A;Residues: 867-880;1148-1217,'Y',1219 <GIM>
A;Experimental source: strain Hall
A;Note: sequence extracted from NCBI backbone (NCBIP:139159); sequence modified after e
R;DasGupta, B.R.; Dekleva, M.L.
Biochimie 72, 661-664, 1990
A;Title: Botulinum neurotoxin type A: sequence of amino acids at the N-terminus and aro
A;Reference number: A60025; MUID:91120847; PMID:2126206
A;Accession: A60025
A;Molecule type: protein
A;Residues: 2-6;445-453,'X',455-457 <DAS1>
R;DasGupta, B.R.; Foley, J.; Niece, R.
Biochemistry 26, 4162, 1987
A;Title: Partial sequence of the light chain of botulinum neurotoxin type A.
A;Reference number: A27000
A;Accession: A27000
A;Molecule type: protein
A;Residues: 2-47 <DAS2>
R;Binz, T.; Blasi, J.; Yamasaki, S.; Baumeister, A.; Link, E.; Suedhof, T.C.; Jahn, R.;
J. Biol. Chem. 269, 1617-1620, 1994
A;Title: Proteolysis of SNAP-25 by types E and A botulinal neurotoxins.
A;Reference number: A49708; MUID:94124495; PMID:8294407
A;Contents: annotation
C;Comment: Botulinum neurotoxins inhibit neurotransmitter release from cholinergic syna
C;Genetics:
A;Gene: atx; botA
C;Function:
A;Description: catalyzes hydrolysis of an Asn-Arg peptide bond in synaptosomal-associat
C;Superfamily: tetanus toxin
C;Keywords: disulfide bond; hydrolase; metalloproteinase; neurotoxin; transmembrane pro
F;2-444/Product: bontoxilysin A light chain #status experimental <LGHT>
F;445-1296/Product: bontoxilysin A heavy chain #status experimental <HVY>
F;223,227/Binding site: zinc (His) #status predicted
F;224/Active site: Glu #status predicted

Query Match 32.5%; Score 65; DB 1; Length 1296;
Best Local Similarity 56.2%; Pred. No. 1;
Matches 9; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

```
Qy          15 SLFNNFTVSFWLRVPK 30
               |:: ||: |||:|:||
Db         936 SMYENFSTSFWIRIPK 951
```

RESULT 4
S33411
botulinum neurotoxin type F - Clostridium barati
C;Species: Clostridium barati
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 16-Jul-1999
C;Accession: S33411; S31860
R;Thompson, D.E.; Hutson, R.A.; East, A.K.; Allaway, D.; Collins, M.D.; Richardson, P.T.
FEMS Microbiol. Lett. 108, 175-182, 1993
A;Title: Nucleotide sequence of the gene coding for Clostridium barati type F neurotoxin
A;Reference number: S33411; MUID:93252228; PMID:8486245
A;Accession: S33411
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1268 <THO>
A;Cross-references: EMBL:X68262; NID:g49138; PIDN:CAA48329.1; PID:g49139
C;Superfamily: tetanus toxin
C;Keywords: neurotoxin

```
  Query Match             32.0%;  Score 64;  DB 2;  Length 1268;
  Best Local Similarity   62.5%;  Pred. No. 1.4;
  Matches   10;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          15 SLFNNFTVSFWLRVPK 30
               | : ||:||||:|:||
Db         920 SRYQNFSVSFWVRIPK 935
```

RESULT 5
I40631
non-proteolytic botulinum neurotoxin type B precursor - Clostridium botulinum
C;Species: Clostridium botulinum
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 16-Jul-1999
C;Accession: I40631; S48103; S48104; S36015
R;Hutson, R.A.; Collins, M.D.; East, A.K.; Thompson, D.E.
Curr. Microbiol. 28, 101-110, 1994
A;Title: Nucleotide sequence of the gene coding for non-proteolytic Clostridium botulinu
A;Reference number: I40631; MUID:94122659; PMID:7764370
A;Accession: I40631
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1291 <RES>
A;Cross-references: EMBL:X71343; NID:g296148; PIDN:CAA50482.1; PID:g296149
R;Campbell, K.D.; Collins, M.D.; East, A.K.
J. Clin. Microbiol. 31, 2255-2262, 1993
A;Title: Gene probes for identification of the botulinal neurotoxin gene and specific id
A;Reference number: S48103; MUID:94013372; PMID:8408542
A;Accession: S48103
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 634-761,'E',763-841,'M',843,'T',845,'N',847-994 <CAM1>
A;Cross-references: EMBL:X70814; NID:g407778; PIDN:CAA50145.1; PID:g407779
A;Experimental source: non-proteolytic strain 2129B (Scott)
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993
A;Accession: S48104
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 634-843,'T',845,'N',847-994 <CAM2>
A;Cross-references: EMBL:X70819; NID:g407780; PIDN:CAA50150.1; PID:g407781
A;Experimental source: non-proteolytic strain Eklund 2B (Colworth 229)
C;Comment: Botulinum neurotoxin type B in these strains may posses a capable catalytic s
C;Genetics:
A;Gene: bont/b
C;Superfamily: tetanus toxin
C;Keywords: metalloprotein; neurotoxin; transmembrane protein; zinc
F;2-441/Product: botulinum neurotoxin type B light chain #status predicted <LGHT>
F;442-1291/Product: botulinum neurotoxin type B heavy chain #status predicted <HVY>
F;230,234/Binding site: zinc (His) #status predicted
F;231/Active site: Glu #status predicted

```
  Query Match             32.0%;  Score 64;  DB 2;  Length 1291;
  Best Local Similarity   62.5%;  Pred. No. 1.4;
  Matches   10;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          15 SLFNNFTVSFWLRVPK 30
               |:| :|:||||:|:||
Db         921 SMFLDFSVSFWIRIPK 936
```

RESULT 6
S48109
neurotoxin type F - Clostridium botulinum (fragment)
C;Species: Clostridium botulinum
C;Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 16-Jul-1999
C;Accession: S48109
R;Campbell, K.D.; Collins, M.D.; East, A.K.
J. Clin. Microbiol. 31, 2255-2262, 1993
A;Title: Gene probes for identification of the botulinal neurotoxin gene and specific
A;Reference number: S48103; MUID:94013372; PMID:8408542
A;Accession: S48109
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-369 <CAM>
A;Cross-references: EMBL:X70820; NID:g407790; PIDN:CAA50151.1; PID:g407791
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993
C;Superfamily: tetanus toxin

```
  Query Match             31.5%;  Score 63;  DB 2;  Length 369;
  Best Local Similarity   56.2%;  Pred. No. 0.48;
  Matches    9;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          15 SLFNNFTVSFWLRVPK 30
               | : ||::|||:|:||
Db         295 SRYQNFSISFWVRIPK 310
```

RESULT 7
I40813
neurotoxin type F - Clostridium botulinum
C;Species: Clostridium botulinum
C;Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 16-Jul-1999
C;Accession: I40813; S48108
R;East, A.K.; Richardson, P.T.; Allaway, D.; Collins, M.D.; Roberts, T.A.; Thompson, D
FEMS Microbiol. Lett. 96, 225-230, 1992
A;Title: Sequence of the gene encoding type F neurotoxin of Clostridium botulinum.
A;Reference number: I40644
A;Accession: I40813
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1274 <RES>
A;Cross-references: GB:M92906; NID:g144866; PIDN:AAA23263.1; PID:g144867
R;Campbell, K.D.; Collins, M.D.; East, A.K.
J. Clin. Microbiol. 31, 2255-2262, 1993
A;Title: Gene probes for identification of the botulinal neurotoxin gene and specific
A;Reference number: S48103; MUID:94013372; PMID:8408542
A;Accession: S48108
A;Status: preliminary; translation not shown
A;Molecule type: DNA
A;Residues: 634-1002 <CAM>
A;Cross-references: EMBL:X70816; NID:g407788; PIDN:CAA50147.1; PID:g407789
C;Superfamily: tetanus toxin
C;Keywords: neurotoxin

```
  Query Match             31.5%;  Score 63;  DB 2;  Length 1274;
  Best Local Similarity   56.2%;  Pred. No. 2;
  Matches    9;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          15 SLFNNFTVSFWLRVPK 30
               | : ||::|||:|:||
Db         928 SRYQNFSISFWVRIPK 943
```

RESULT 8
A48940
bontoxilysin (EC 3.4.24.69) B precursor - Clostridium botulinum
N;Alternate names: botulinum neurotoxin type B (BoNT/B)
C;Species: Clostridium botulinum
C;Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 18-Jun-1999
C;Accession: A48940; S48105; S21575; A42871; S07155; S08562; S07128; S08573; S08574
R;Whelan, S.M.; Elmore, M.J.; Bodsworth, N.J.; Brehm, J.K.; Atkinson, T.; Minton, N.P.
Appl. Environ. Microbiol. 58, 2345-2354, 1992
A;Title: Molecular cloning of the Clostridium botulinum structural gene encoding the typ
A;Reference number: A48940; MUID:92384550; PMID:1514783
A;Accession: A48940
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1291 <WHE>
A;Cross-references: GB:M81186; NID:g144734; PIDN:AAA23211.1; PID:g144735
A;Experimental source: type B, Danish
A;Note: sequence extracted from NCBI backbone (NCBIN:112080, NCBIP:112081); this publica
R;Campbell, K.D.; Collins, M.D.; East, A.K.
J. Clin. Microbiol. 31, 2255-2262, 1993
A;Title: Gene probes for identification of the botulinal neurotoxin gene and specific id
A;Reference number: S48103; MUID:94013372; PMID:8408542
A;Accession: S48105
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 634-994 <CAM>
A;Cross-references: EMBL:X70817; NID:g407782; PIDN:CAA50148.1; PID:g407783
A;Experimental source: proteolytic type B, strain NCTC 7273
R;Szabo, E.A.; Pemberton, J.M.; Desmarchelier, P.M.
submitted to the EMBL Data Library, April 1992
A;Description: Partial amino acid sequence of botulinum neurotoxin type B and comparisio
A;Reference number: S21575
A;Accession: S21575
A;Molecule type: DNA
A;Residues: 36-217,'G',219-224,'S',226-246 <SZA>
A;Cross-references: EMBL:Z11934; NID:g40383; PIDN:CAA77991.1; PID:g40384
R;Kurazono, H.; Mochida, S.; Binz, T.; Eisel, U.; Quanz, M.; Grebenstein, O.; Wernars, K
J. Biol. Chem. 267, 14721-14729, 1992
A;Title: Minimal essential domains specifying toxicity of the light chains of tetanus to
A;Reference number: A42871; MUID:92340509; PMID:1634516
A;Accession: A42871
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-313,'S',315-451 <KUR>
A;Experimental source: strain Okra
A;Note: sequence extracted from NCBI backbone (NCBIP:109365)
R;DasGupta, B.R.; Datta, A.
Biochimie 70, 811-817, 1988
A;Title: Botulinum neurotoxin type B (strain 657): partial sequence and similarity with
A;Reference number: S07155; MUID:89000987; PMID:3139097
A;Accession: S07155
A;Molecule type: protein
A;Residues: 2-29,'M',31-45 <DAS>
A;Accession: S08562
A;Molecule type: protein
A;Residues: 442-463,'R',465-467 <DA2>
R;Schmidt, J.J.; Sathyamoorthy, V.; DasGupta, B.R.
Arch. Biochem. Biophys. 238, 544-548, 1985
A;Title: Partial amino acid sequences of botulinum neurotoxins types B and E.
A;Reference number: S07128; MUID:85197963; PMID:3888113
A;Accession: S07128
A;Status: preliminary
A;Molecule type: protein
A;Residues: 2-16 <SCH1>
A;Accession: S08573
A;Status: preliminary
A;Molecule type: protein
A;Residues: 2-17 <SCH2>
A;Accession: S08574
A;Status: preliminary
A;Molecule type: protein
A;Residues: 442-459 <SCH3>
R;Schiavo, G.; Benfenati, F.; Poulain, B.; Rossetto, O.; de Laureto, P.P.; DasGupta, B.
Nature 359, 832-835, 1992
A;Title: Tetanus and botulinum-B neurotoxins block neurotransmitter release by proteoly
A;Reference number: S27125; MUID:93063293; PMID:1331807
A;Contents: annotation
C;Comment: Botulinum neurotoxins inhibit neurotransmitter release from cholinergic syna
C;Genetics:
A;Gene: bont/b
C;Function:
A;Description: catalyzes hydrolysis of a Gln-Phe peptide bond in synaptobrevin 2
C;Superfamily: tetanus toxin
C;Keywords: hydrolase; metalloproteinase; neurotoxin; transmembrane protein; zinc
F;2-441/Product: bontoxilysin B light chain #status experimental <LGHT>
F;442-1291/Product: bontoxilysin B heavy chain #status experimental <HVY>
F;230,234/Binding site: zinc (His) #status predicted
F;231/Active site: Glu #status predicted

  Query Match 31.5%; Score 63; DB 1; Length 1291;
  Best Local Similarity 62.5%; Pred. No. 2;
  Matches 10; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 15 SLFNNFTVSFWLRVPK 30
Db 921 SVFLDFSVSFWIRIPK 936

RESULT 9
I40645
botulinum neurotoxin type A - Clostridium botulinum
C;Species: Clostridium botulinum
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 16-Jul-1999
C;Accession: I40645
R;Willems, A.; East, A.K.; Lawson, P.A.; Collins, M.D.
Res. Microbiol. 144, 547-556, 1993
A;Title: Sequence of the gene coding for the neurotoxin of Clostridium botulinum type A
A;Reference number: I40645; MUID:94143603; PMID:8310180
A;Accession: I40645
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1296 <RES>
A;Cross-references: EMBL:X73423; NID:g507070; PIDN:CAA51824.1; PID:g507071
C;Superfamily: tetanus toxin
C;Keywords: neurotoxin

  Query Match 31.0%; Score 62; DB 2; Length 1296;
  Best Local Similarity 50.0%; Pred. No. 2.8;
  Matches 8; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

Qy 15 SLFNNFTVSFWLRVPK 30
Db 936 SMYENFSTSFWIKIPK 951

RESULT 10
S48110
neurotoxin type F - Clostridium botulinum (fragment)
C;Species: Clostridium botulinum
C;Date: 14-Jul-1995 #sequence_revision 10-Nov-1995 #text_change 16-Jul-1999
C;Accession: S48110
R;Campbell, K.D.; Collins, M.D.; East, A.K.
J. Clin. Microbiol. 31, 2255-2262, 1993
A;Title: Gene probes for identification of the botulinal neurotoxin gene and specific i
A;Reference number: S48103; MUID:94013372; PMID:8408542
A;Accession: S48110
A;Status: preliminary; translation not shown
A;Molecule type: DNA
A;Residues: 1-366 <CAM>
A;Cross-references: EMBL:X70821; NID:g407792; PIDN:CAA50152.1; PID:g407793
C;Superfamily: tetanus toxin
C;Keywords: neurotoxin

  Query Match 30.5%; Score 61; DB 2; Length 366;
  Best Local Similarity 57.1%; Pred. No. 0.91;

```
  Matches    8;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         17 FNNPTVSFWLRVPK 30
              : ||::|||:|:||
Db        297 YQNPSISFWVRIPK 310


RESULT 11
I78541
gonadoliberin precursor - rhesus macaque (fragment)
N;Alternate names: luteinizing hormone releasing hormone
C;Species: Macaca mulatta (rhesus macaque)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 16-Jul-1999
C;Accession: I78541
R;Ma, Y.J.; Costa, M.E.; Ojeda, S.R.
Neuroendocrinology 60, 346-359, 1994
A;Title: Developmental expression of the genes encoding transforming growth factor alpha
A;Reference number: I58134; MUID:95124501; PMID:7545971
A;Accession: I78541
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-67 <RES>
A;Cross-references: GB:S75918; NID:g912831; PIDN:AAB33096.1; PID:g912832
C;Superfamily: gonadoliberin

  Query Match             29.5%;  Score 59;  DB 2;  Length 67;
  Best Local Similarity   52.4%;  Pred. No. 0.26;
  Matches   11;  Conservative    3;  Mismatches    5;  Indels    2;  Gaps    1;

Qy          2 HWSYGLRPGS--SGPSLFNNF 20
              |||||||||      :| ::|
Db          7 HWSYGLRPGGKRDAENLMDSF 27


RESULT 12
RHHUG
gonadoliberin precursor [validated] - human
N;Alternate names: gonadotropin releasing hormone (GnRH); luteinizing hormone releasing
N;Contains: gonadoliberin-associated protein (GAP); progonadoliberin
C;Species: Homo sapiens (man)
C;Date: 17-Mar-1987 #sequence_revision 21-Jul-1995 #text_change 08-Dec-2000
C;Accession: S05308; A26173; A93342; A90108; A01410; S45718
R;Hayflick, J.S.; Adelman, J.P.; Seeburg, P.H.
Nucleic Acids Res. 17, 6403-6404, 1989
A;Title: The complete nucleotide sequence of the human gonadotropin-releasing hormone ge
A;Reference number: S05308; MUID:89366682; PMID:2671939
A;Accession: S05308
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-92 <HAY>
A;Cross-references: EMBL:X15215; NID:g31955; PIDN:CAA33285.1; PID:g31956
R;Adelman, J.P.; Mason, A.J.; Hayflick, J.S.; Seeburg, P.H.
Proc. Natl. Acad. Sci. U.S.A. 83, 179-183, 1986
A;Title: Isolation of the gene and hypothalamic cDNA for the common precursor of gonadot
A;Reference number: A94090; MUID:86094338; PMID:2867548
A;Accession: A26173
A;Molecule type: mRNA
A;Residues: 1-92 <ADE>
A;Cross-references: GB:M12578; NID:g183418; PIDN:AAA35916.1; PID:g386749
A;Experimental source: hypothalamus
R;Seeburg, P.H.; Adelman, J.P.
Nature 311, 666-668, 1984
A;Title: Characterization of cDNA for precursor of human luteinizing hormone releasing h
A;Reference number: A93342; MUID:85012739; PMID:6090951
A;Accession: A93342
A;Molecule type: mRNA
A;Residues: 1-15,'S',17-92 <SEE>
A;Cross-references: GB:X01059; NID:g34356; PIDN:CAA25526.1; PID:g34357
A;Experimental source: placenta
R;Tan, L.; Rousseau, P.
Biochem. Biophys. Res. Commun. 109, 1061-1071, 1982
A;Title: The chemical identity of the immunoreactive LHRH-like peptide biosynthesized in
A;Reference number: A90108; MUID:83126573; PMID:6760865
A;Accession: A90108
A;Molecule type: protein
A;Residues: 24-33 <TAN>
A;Experimental source: placental trophoblasts
R;Leibovitz, D.; Koch, Y.; Pitzer, F.; Fridkin, M.; Dantes, A.; Baumeister, W.; Amster
FEBS Lett. 346, 203-206, 1994
A;Title: Sequential degradation of the neuropeptide gonadotropin-releasing hormone by
A;Reference number: S45718; MUID:94283597; PMID:8013634
A;Contents: annotation; degradation pathway of synthetic hormone
C;Genetics:
A;Gene: GDB:GNRH; LHRH; GRH
A;Cross-references: GDB:133746; OMIM:227200; OMIM:152760
A;Map position: 8p21-8p11.2
A;Introns: 47/3; 79/3
C;Function:
A;Description: gonadoliberin stimulates pituitary secretion of lutropin and follitropi
A;Note: gonadoliberin-associated protein may have prolactin release inhibiting activiy
C;Superfamily: gonadoliberin
C;Keywords: amidated carboxyl end; hormone; hypothalamus; placenta; pyroglutamic acid
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-92/Product: progonadoliberin #status predicted <PGN>
F;24-33/Product: gonadoliberin #status experimental <MAT>
F;37-92/Product: gonadoliberin-associated protein #status predicted <GAP>
F;24/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experim
F;33/Modified site: amidated carboxyl end (Gly) (amide in mature form from following g

  Query Match             29.5%;  Score 59;  DB 1;  Length 92;
  Best Local Similarity   52.4%;  Pred. No. 0.37;
  Matches   11;  Conservative    3;  Mismatches    5;  Indels    2;  Gaps    1;

Qy          2 HWSYGLRPGS--SGPSLFNNF 20
              |||||||||      :| ::|
Db         25 HWSYGLRPGGKRDAENLIDSF 45


RESULT 13
RHPGG
gonadoliberin - pig
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 18-Mar-1997
C;Accession: A01411
R;Baba, Y.; Matsuo, H.; Schally, A.V.
Biochem. Biophys. Res. Commun. 44, 459-463, 1971
A;Title: Structure of the porcine LH- and FSH-releasing hormone. II. Confirmation of t
A;Reference number: A90172; MUID:72114303; PMID:4946067
A;Accession: A01411
A;Molecule type: protein
A;Residues: 1-10 <BAB>
R;Matsuo, H.; Arimura, A.; Nair, R.M.G.; Schally, A.V.
Biochem. Biophys. Res. Commun. 45, 822-827, 1971
A;Title: Synthesis of the porcine LH- and FSH-releasing hormone by the solid-phase met
A;Reference number: A90176; MUID:72065376; PMID:4942726
A;Contents: annotation; synthesis
A;Note: the synthetic and natural hormones have the same physicochemical and biologica
R;Baba, Y.; Arimura, A.; Schally, A.V.
Biochem. Biophys. Res. Commun. 45, 483-487, 1971
A;Title: On the tryptophan residue in porcine LH and FSH-releasing hormone.
A;Reference number: A90175; MUID:72117544; PMID:4946275
A;Contents: annotation
A;Note: Trp-3 appears to be essential for biological activity
C;Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and
C;Superfamily: gonadoliberin
C;Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;10/Modified site: amidated carboxyl end (Gly) #status experimental

  Query Match             29.0%;  Score 58;  DB 1;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 0.041;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 HWSYGLRPG 10
              |||||||||
Db          2 HWSYGLRPG 10
```

RESULT 14
RHSHG
gonadoliberin - sheep
C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Mar-1997
C;Accession: A93780; A01411
R;Burgus, R.; Butcher, M.; Amoss, M.; Ling, N.; Monahan, M.; Rivier, J.; Fellows, R.; Bl
Proc. Natl. Acad. Sci. U.S.A. 69, 278-282, 1972
A;Title: Primary structure of the ovine hypothalamic luteinizing hormone-releasing facto
A;Reference number: A93780; MUID:72094314; PMID:4550508
A;Accession: A93780
A;Molecule type: protein
A;Residues: 1-10 <BUR>
A;Note: the natural and synthetic hormones have the same biological activity
C;Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and fo
C;Superfamily: gonadoliberin
C;Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 29.0%; Score 58; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.041;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          2 HWSYGLRPG 10
              |||||||||
Db          2 HWSYGLRPG 10
```

RESULT 15
I51423
gonadoliberin precursor - African clawed frog
N;Alternate names: luteinizing hormone releasing hormone
C;Species: Xenopus laevis (African clawed frog)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 16-Jul-1999
C;Accession: I51423
R;Hayes, W.P.; Wray, S.; Battey, J.F.
Endocrinology 134, 1835-1845, 1994
A;Title: The frog GnRH-I gene has a mammalian-like expression pattern and conserved doma
A;Reference number: I51423; MUID:94185563; PMID:8137750
A;Accession: I51423
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-89 <HAY>
A;Cross-references: GB:L28040; NID:g496291; PIDN:AAA49728.1; PID:g496292
C;Genetics:
A;Gene: GnRH-I
C;Superfamily: gonadoliberin

Query Match 29.0%; Score 58; DB 2; Length 89;
Best Local Similarity 100.0%; Pred. No. 0.49;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          2 HWSYGLRPG 10
              |||||||||
Db         25 HWSYGLRPG 33
```

Search completed: March 10, 2004, 09:16:50
Job time : 12.5763 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:54 ; Search time 36.5681 Seconds
(without alignments)
319.245 Million cell updates/sec

Title: US-09-848-834A-14
Perfect score: 200
Sequence: 1 XHWSYGLRPGSSGPSLFNNFTVSFWLRVPKVSASHLE 37

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	114	57.0	451	2	Q9LA13	Q9la13 clostridium
2	114	57.0	1310	2	Q93N27	Q93n27 clostridium
3	64	32.0	361	2	Q45848	Q45848 clostridium
4	64	32.0	361	2	Q45846	Q45846 clostridium
5	64	32.0	441	2	Q9X708	Q9x708 clostridium
6	64	32.0	1268	2	Q45851	Q45851 clostridium
7	64	32.0	1291	2	Q9ZAJ8	Q9zaj8 clostridium
8	64	32.0	1291	2	Q93G71	Q93g71 clostridium
9	64	32.0	1291	2	Q933K0	Q933k0 clostridium
10	64	32.0	1291	2	Q08077	Q08077 clostridium
11	64	32.0	1291	2	Q8GR96	Q8gr96 clostridium
12	61	30.5	1278	2	Q57236	Q57236 clostridium
13	58	29.0	91	13	Q9PRH0	Q9prh0 anguilla ja
14	57.5	28.7	234	12	Q9W9A5	Q9w9a5 simian viru
15	57.5	28.7	234	12	O92837	O92837 simian viru
16	57.5	28.7	352	12	Q98VM1	Q98vm1 simian viru
17	57.5	28.7	352	12	Q9W9F7	Q9w9f7 simian viru
18	57.5	28.7	352	12	Q80FH1	Q80fh1 simian viru
19	57.5	28.7	352	12	Q910V5	Q910v5 simian viru
20	57	28.5	1280	2	Q9ZAJ5	Q9zaj5 clostridium
21	56.5	28.2	94	13	Q8JFY3	Q8jfy3 oreochromis
22	56.5	28.2	98	13	Q805A5	Q805a5 oreochromis
23	56	28.0	367	2	Q45861	Q45861 clostridium
24	56	28.0	367	2	Q45862	Q45862 clostridium
25	56	28.0	1251	2	Q9K395	Q9k395 clostridium
26	56	28.0	1252	2	Q8KZM3	Q8kzm3 clostridium
27	56	28.0	1255	2	Q9FAR6	Q9far6 clostridium
28	56	28.0	1291	2	Q93HT3	Q93ht3 clostridium
29	54.5	27.3	467	12	Q8JXK2	Q8jxk2 tomato spot
30	54.5	27.3	467	12	Q8JVL0	Q8jvl0 tomato spot
31	54	27.0	200	4	Q86YU6	Q86yu6 homo sapien
32	53.5	26.8	251	16	Q92AM6	Q92am6 listeria in
33	53.5	26.8	609	12	Q8B6X9	Q8b6x9 peste-des-p
34	53	26.5	493	16	Q7VLI1	Q7vli1 haemophilus
35	53	26.5	1196	2	Q45916	Q45916 clostridium
36	53	26.5	1196	2	Q53550	Q53550 clostridium
37	53	26.5	1196	2	Q9LBR2	Q9lbr2 clostridium
38	53	26.5	1196	2	Q9LBS8	Q9lbs8 clostridium
39	53	26.5	1196	2	Q93HT4	Q93ht4 clostridium
40	53	26.5	1196	9	Q9ZX77	Q9zx77 clostridium
41	53	26.5	1196	9	Q38197	Q38197 clostridium
42	53	26.5	1285	2	Q45967	Q45967 clostridium
43	53	26.5	1285	2	Q9LBR1	Q9lbr1 clostridium
44	53	26.5	1702	12	Q8JXI5	Q8jxi5 norwalk-lik
45	53	26.5	1702	12	Q8JXI4	Q8jxi4 norwalk-lik

ALIGNMENTS

RESULT 1
Q9LA13
ID Q9LA13 PRELIMINARY; PRT; 451 AA.
AC Q9LA13;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Tetanus toxin (Fragment).
OS Clostridium tetani.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1513;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=20886;
RA He H.J., Shi H.J., He Z.Y., Yuan Q.S., Wu X.F.;
RT "Fragment C of Tetanus Toxin.";
RL Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF154828; AAF73267.1; -.
DR HSSP; P04958; 1A8D.
DR GO; GO:0004866; F:endopeptidase inhibitor activity; IEA.
DR InterPro; IPR008985; ConA_like_lec_gl.
DR InterPro; IPR001064; Crystallin.
DR InterPro; IPR002160; Kunitz_legume.
DR PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; 1.
FT NON_TER 1 1
SQ SEQUENCE 451 AA; 51823 MW; 69A8C5F030E6CD8E CRC64;

Query Match 57.0%; Score 114; DB 2; Length 451;
Best Local Similarity 95.5%; Pred. No. 1e-07;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          16 LFNNFTVSFWLRVPKVSASHLE 37
               :||||||||||||||||||||||
Db          82 MFNNFTVSFWLRVPKVSASHLE 103
```

RESULT 2
Q93N27

```
ID   Q93N27      PRELIMINARY;      PRT;  1310 AA.
AC   Q93N27;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Tetanus toxin (Fragment).
OS   Clostridium tetani.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1513;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Shumin Z., Dianliang L.;
RT   "Cloning and sequence analysis of tetanus toxin gene.";
RL   Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF389424; AAK72964.2; -.
DR   GO; GO:0004866; F:endopeptidase inhibitor activity; IEA.
DR   GO; GO:0008237; F:metallopeptidase activity; IEA.
DR   GO; GO:0015070; F:toxin activity; IEA.
DR   GO; GO:0008270; F:zinc ion binding; IEA.
DR   GO; GO:0009405; P:pathogenesis; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR008985; ConA_like_lec_gl.
DR   InterPro; IPR001064; Crystallin.
DR   InterPro; IPR002160; Kunitz_legume.
DR   InterPro; IPR000395; Peptidase_M27.
DR   InterPro; IPR006025; Pept_M_Zn_BS.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
FT   NON_TER       1      1
FT   NON_TER    1310   1310
SQ   SEQUENCE   1310 AA;  150316 MW;  9EADDC914418E450 CRC64;

  Query Match             57.0%;  Score 114;  DB 2;  Length 1310;
  Best Local Similarity   95.5%;  Pred. No. 3.2e-07;
  Matches   21;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         16 LFNNFTVSFWLRVPKVSASHLE 37
              :||||||||||||||||||||||
Db        947 MFNNFTVSFWLRVPKVSASHLE 968


RESULT 3
Q45848
ID   Q45848      PRELIMINARY;      PRT;   361 AA.
AC   Q45848;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Botulinum neurotoxin type B (Fragment).
GN   BONT/B.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=type B;
RX   MEDLINE=94013372; PubMed=8408542;
RA   Campbell K., East A.K., Collins M.D.;
RT   "Gene probes for identification of the botulinal neurotoxin gene and
RT   specific identification of neurotoxin types B, E, and F.";
RL   J. Clin. Microbiol. 31:2255-2262(1993).
DR   EMBL; X70819; CAA50150.1; -.
DR   HSSP; P10845; 3BTA.
DR   GO; GO:0015070; F:toxin activity; IEA.
DR   InterPro; IPR008985; ConA_like_lec_gl.
KW   Neurotoxin.
FT   NON_TER       1      1
FT   NON_TER     361    361
SQ   SEQUENCE   361 AA;  42131 MW;  A2E0FFFC81F9533D CRC64;

  Query Match             32.0%;  Score 64;  DB 2;  Length 361;
  Best Local Similarity   62.5%;  Pred. No. 1;
  Matches   10;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         15 SLFNNFTVSFWLRVPK 30
              |:| :|:||||:|:||
Db        288 SMFLDFSVSFWIRIPK 303


RESULT 4
Q45846
ID   Q45846      PRELIMINARY;      PRT;   361 AA.
AC   Q45846;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Botulinum neurotoxin type B (Fragment).
GN   BONT/B.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=type B;
RX   MEDLINE=94013372; PubMed=8408542;
RA   Campbell K., East A.K., Collins M.D.;
RT   "Gene probes for identification of the botulinal neurotoxin gene and
RT   specific identification of neurotoxin types B, E, and F.";
RL   J. Clin. Microbiol. 31:2255-2262(1993).
DR   EMBL; X70814; CAA50145.1; -.
DR   HSSP; P10845; 3BTA.
DR   GO; GO:0015070; F:toxin activity; IEA.
DR   InterPro; IPR008985; ConA_like_lec_gl.
KW   Neurotoxin.
FT   NON_TER       1      1
FT   NON_TER     361    361
SQ   SEQUENCE   361 AA;  42175 MW;  533EA98735CD98E1 CRC64;

  Query Match             32.0%;  Score 64;  DB 2;  Length 361;
  Best Local Similarity   62.5%;  Pred. No. 1;
  Matches   10;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         15 SLFNNFTVSFWLRVPK 30
              |:| :|:||||:|:||
Db        288 SMFLDFSVSFWIRIPK 303


RESULT 5
Q9X708
ID   Q9X708      PRELIMINARY;      PRT;   441 AA.
AC   Q9X708;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Botulinum neurotoxin type B (Fragment).
GN   BONT/B.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99343691; PubMed=10413679;
RA   Lalli G., Herreros J., Osborne S.L., Montecucco C., Rossetto O.,
RA   Schiavo G.;
RT   "Functional characterisation of tetanus and botulinum neurotoxins
RT   binding domains.";
RL   J. Cell Sci. 112:2715-2724(1999).
DR   EMBL; AJ242628; CAB43706.1; -.
DR   HSSP; P10845; 3BTA.
```

```
DR   GO; GO:0004866; F:endopeptidase inhibitor activity; IEA.
DR   GO; GO:0015070; F:toxin activity; IEA.
DR   InterPro; IPR008985; ConA_like_lec_gl.
DR   InterPro; IPR002160; Kunitz_legume.
KW   Neurotoxin.
FT   NON_TER       1      1
FT   NON_TER     441    441
SQ   SEQUENCE   441 AA;  52772 MW;  721D0B468E8C95A4 CRC64;

  Query Match             32.0%;  Score 64;  DB 2;  Length 441;
  Best Local Similarity   62.5%;  Pred. No. 1.3;
  Matches   10;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         15 SLFNNFTVSFWLRVPK 30
              |:| :|:||||:|:||
Db         79 SMFLDFSVSFWIRIPK 94


RESULT 6
Q45851
ID   Q45851      PRELIMINARY;      PRT;  1268 AA.
AC   Q45851;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Neurotoxin type F.
GN   BONT /F.
OS   Clostridium baratii.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1561;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93252228; PubMed=8486245;
RA   Thompson D.E., Hutson R.A., East A.K., Allaway D., Collins M.D.,
RA   Richardson P.T.;
RT   "Nucleotide sequence of the gene coding for Clostridium barati type F
RT   neurotoxin: Comparison with other clostridial neurotoxins.";
RL   FEMS Microbiol. Lett. 108:175-182(1993).
DR   EMBL; X68262; CAA48329.1; -.
DR   PIR; S33411; S33411.
DR   HSSP; P10845; 3BTA.
DR   MEROPS; M27.002; -.
DR   GO; GO:0004866; F:endopeptidase inhibitor activity; IEA.
DR   GO; GO:0008237; F:metallopeptidase activity; IEA.
DR   GO; GO:0015070; F:toxin activity; IEA.
DR   GO; GO:0008270; F:zinc ion binding; IEA.
DR   GO; GO:0009405; P:pathogenesis; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR008985; ConA_like_lec_gl.
DR   InterPro; IPR002160; Kunitz_legume.
DR   InterPro; IPR000395; Peptidase_M27.
DR   InterPro; IPR006025; Pept_M_Zn_BS.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
SQ   SEQUENCE   1268 AA;  145513 MW;  963040091AC15ED2 CRC64;

  Query Match             32.0%;  Score 64;  DB 2;  Length 1268;
  Best Local Similarity   62.5%;  Pred. No. 4;
  Matches   10;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         15 SLFNNFTVSFWLRVPK 30
              | : ||:||||:|:||
Db        920 SRYQNFSVSFWVRIPK 935


RESULT 7
Q9ZAJ8
ID   Q9ZAJ8      PRELIMINARY;      PRT;  1291 AA.
AC   Q9ZAJ8;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   BonT protein.
GN   BONT.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CDC 3281;
RX   MEDLINE=98440323; PubMed=9767710;
RA   Santos-Buelga J., Collins M.D., East A.K.;
RT   "Characterization of the genes encoding the Botulinum neurotoxin
RT   complex in a strain of clostridium botulinum producing type B & F
RT   neurotoxins.";
RL   Curr. Microbiol. 37:312-318(1998).
DR   EMBL; Y13630; CAA73968.1; -.
DR   HSSP; P10845; 3BTA.
DR   GO; GO:0004866; F:endopeptidase inhibitor activity; IEA.
DR   GO; GO:0008237; F:metallopeptidase activity; IEA.
DR   GO; GO:0015070; F:toxin activity; IEA.
DR   GO; GO:0008270; F:zinc ion binding; IEA.
DR   GO; GO:0009405; P:pathogenesis; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR008985; ConA_like_lec_gl.
DR   InterPro; IPR002160; Kunitz_legume.
DR   InterPro; IPR000395; Peptidase_M27.
DR   InterPro; IPR006025; Pept_M_Zn_BS.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
SQ   SEQUENCE   1291 AA;  150840 MW;  E4D3B0E46AB2E735 CRC64;

  Query Match             32.0%;  Score 64;  DB 2;  Length 1291;
  Best Local Similarity   62.5%;  Pred. No. 4.1;
  Matches   10;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         15 SLFNNFTVSFWLRVPK 30
              |:| :|:||||:|:||
Db        921 SMFLDFSVSFWIRIPK 936


RESULT 8
Q93G71
ID   Q93G71      PRELIMINARY;      PRT;  1291 AA.
AC   Q93G71;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Neurotoxin type B.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1436;
RA   Kirma N., Ferreira J.L., Baumstark B.R.;
RT   "Characterization of six type A strains of Clostridium botulinum that
RT   contain type B toxin gene sequences.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF295926; AAK97132.1; -.
DR   GO; GO:0004866; F:endopeptidase inhibitor activity; IEA.
DR   GO; GO:0008237; F:metallopeptidase activity; IEA.
DR   GO; GO:0015070; F:toxin activity; IEA.
DR   GO; GO:0008270; F:zinc ion binding; IEA.
DR   GO; GO:0009405; P:pathogenesis; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR008985; ConA_like_lec_gl.
DR   InterPro; IPR002160; Kunitz_legume.
DR   InterPro; IPR000395; Peptidase_M27.
```

```
DR   InterPro; IPR006025; Pept_M_Zn_BS.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
SQ   SEQUENCE   1291 AA;  150824 MW;  D7CA07BAE2EB8CD2 CRC64;

  Query Match             32.0%;  Score 64;  DB 2;  Length 1291;
  Best Local Similarity   62.5%;  Pred. No. 4.1;
  Matches   10;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         15 SLFNNFTVSFWLRVPK 30
              |:| :|:||||:|:||
Db        921 SMFLDFSVSFWIRIPK 936


RESULT 9
Q933K0
ID   Q933K0      PRELIMINARY;      PRT;  1291 AA.
AC   Q933K0;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Type B cryptic neurotoxin.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=593, and 588;
RA   Kirma N., Ferreira J.L., Baumstark B.R.;
RT   "Characterization of six type A strains of Clostridium botulinum that
RT   contain type B toxin gene sequences.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF300466; AAL11499.1; -.
DR   EMBL; AF300465; AAL11498.1; -.
DR   GO; GO:0004866; F:endopeptidase inhibitor activity; IEA.
DR   GO; GO:0008237; F:metallopeptidase activity; IEA.
DR   GO; GO:0015070; F:toxin activity; IEA.
DR   GO; GO:0008270; F:zinc ion binding; IEA.
DR   GO; GO:0009405; P:pathogenesis; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR008985; ConA_like_lec_gl.
DR   InterPro; IPR002160; Kunitz_legume.
DR   InterPro; IPR000395; Peptidase_M27.
DR   InterPro; IPR006025; Pept_M_Zn_BS.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
KW   Neurotoxin.
SQ   SEQUENCE   1291 AA;  150843 MW;  7AC1737B0FA5A151 CRC64;

  Query Match             32.0%;  Score 64;  DB 2;  Length 1291;
  Best Local Similarity   62.5%;  Pred. No. 4.1;
  Matches   10;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         15 SLFNNFTVSFWLRVPK 30
              |:| :|:||||:|:||
Db        921 SMFLDFSVSFWIRIPK 936


RESULT 10
Q08077
ID   Q08077      PRELIMINARY;      PRT;  1291 AA.
AC   Q08077;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   BoNT/B.
GN   BONT/B.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Eklund 17B ATCC25765;
RX   MEDLINE=94122659; PubMed=7764370;
RA   Hutson R.A., Collins M.D., East A.K., Thompson D.E.;
RT   "Nucleotide sequence of the gene coding for non-proteolytic
RT   Clostridium botulinum type B neurotoxin: comparison with other
RT   clostridial neurotoxins.";
RL   Curr. Microbiol. 28:101-110(1994).
DR   EMBL; X71343; CAA50482.1; -.
DR   PIR; I40631; I40631.
DR   HSSP; P10845; 3BTA.
DR   MEROPS; M27.002; -.
DR   GO; GO:0004866; F:endopeptidase inhibitor activity; IEA.
DR   GO; GO:0008237; F:metallopeptidase activity; IEA.
DR   GO; GO:0015070; F:toxin activity; IEA.
DR   GO; GO:0008270; F:zinc ion binding; IEA.
DR   GO; GO:0009405; P:pathogenesis; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR008985; ConA_like_lec_gl.
DR   InterPro; IPR002160; Kunitz_legume.
DR   InterPro; IPR000395; Peptidase_M27.
DR   InterPro; IPR006025; Pept_M_Zn_BS.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
SQ   SEQUENCE   1291 AA;  150513 MW;  71BCAFE23D69FAAA CRC64;

  Query Match             32.0%;  Score 64;  DB 2;  Length 1291;
  Best Local Similarity   62.5%;  Pred. No. 4.1;
  Matches   10;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         15 SLFNNFTVSFWLRVPK 30
              |:| :|:||||:|:||
Db        921 SMFLDFSVSFWIRIPK 936


RESULT 11
Q8GR96
ID   Q8GR96      PRELIMINARY;      PRT;  1291 AA.
AC   Q8GR96;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Neurotoxin.
GN   BONTB.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Ihara H., Kohda T., Morimoto F., Tsukamoto K., Karasawa T.,
RA   Nakamura S., Mukamoto M., Kozaki S.;
RT   "Clostridium botulinum type B neurotoxin associated with infant
RT   botulism.";
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB084152; BAC22064.1; -.
DR   GO; GO:0004866; F:endopeptidase inhibitor activity; IEA.
DR   GO; GO:0008237; F:metallopeptidase activity; IEA.
DR   GO; GO:0015070; F:toxin activity; IEA.
DR   GO; GO:0008270; F:zinc ion binding; IEA.
DR   GO; GO:0009405; P:pathogenesis; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR008985; ConA_like_lec_gl.
DR   InterPro; IPR002160; Kunitz_legume.
DR   InterPro; IPR000395; Peptidase_M27.
DR   InterPro; IPR006025; Pept_M_Zn_BS.
DR   Pfam; PF01742; Peptidase_M27; 1.
```

```
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
SQ   SEQUENCE   1291 AA;  150574 MW;  0227CAEF4F58504D CRC64;

  Query Match             32.0%;  Score 64;  DB 2;  Length 1291;
  Best Local Similarity   62.5%;  Pred. No. 4.1;
  Matches   10;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         15 SLFNNFTVSFWLRVPK 30
              |:| :|:||||:|:||
Db        921 SMFLDFSVSFWIRIPK 936


RESULT 12
Q57236
ID   Q57236      PRELIMINARY;      PRT;  1278 AA.
AC   Q57236; Q45863;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Botulinum neurotoxin type F (BONT/F protein).
GN   BONT/F.
OS   Clostridium botulinum.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NCTC 10281;
RA   Hutson R.A., Collins M.D.;
RL   Submitted (AUG-1995) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Elmore M.J., Bodsworth N.J., Whelan S.M., Minton N.P.;
RL   Submitted (AUG-1994) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 635-1000 FROM N.A.
RC   STRAIN=NCTC 1028;
RX   MEDLINE=94013372; PubMed=8408542;
RA   Campbell K., East A.K., Collins M.D.;
RT   "Gene probes for identification of the botulinal neurotoxin gene and
RT   specific identification of neurotoxin types B, E, and F.";
RL   J. Clin. Microbiol. 31:2255-2262(1993).
RN   [4]
RP   SEQUENCE OF 1-27 FROM N.A.
RC   STRAIN=LANGELAND;
RX   MEDLINE=98404102; PubMed=9732534;
RA   East A.K., Bhandari M., Hielm S., Collins M.D.;
RT   "Analysis of the botulinum neurotoxin type F gene clusters in
RT   proteolytic and nonproteolytic Clostridium botulinum and Clostridium
RT   barati.";
RL   Curr. Microbiol. 37:262-268(1998).
DR   EMBL; X81714; CAA57358.1; -.
DR   EMBL; L35496; AAA23210.1; -.
DR   EMBL; X70821; CAA50152.1; -.
DR   EMBL; X99064; CAA67512.1; -.
DR   PIR; S48110; S48110.
DR   HSSP; P10845; 3BTA.
DR   MEROPS; M27.002; -.
DR   GO; GO:0004866; F:endopeptidase inhibitor activity; IEA.
DR   GO; GO:0008237; F:metallopeptidase activity; IEA.
DR   GO; GO:0015070; F:toxin activity; IEA.
DR   GO; GO:0008270; F:zinc ion binding; IEA.
DR   GO; GO:0009405; P:pathogenesis; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR008985; ConA_like_lec_gl.
DR   InterPro; IPR002160; Kunitz_legume.
DR   InterPro; IPR000395; Peptidase_M27.
DR   InterPro; IPR006025; Pept_M_Zn_BS.
DR   Pfam; PF01742; Peptidase_M27; 1.
DR   PRINTS; PR00760; BONTOXILYSIN.
DR   ProDom; PD001963; Bontoxilysin; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
KW   Neurotoxin.
SQ   SEQUENCE   1278 AA;  147073 MW;  A1BE1318431D6918 CRC64;

  Query Match             30.5%;  Score 61;  DB 2;  Length 1278;
  Best Local Similarity   57.1%;  Pred. No. 11;
  Matches    8;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         17 FNNFTVSFWLRVPK 30
              : ||::|||:|:||
Db        931 YQNFSISFWVRIPK 944


RESULT 13
Q9PRH0
ID   Q9PRH0      PRELIMINARY;      PRT;    91 AA.
AC   Q9PRH0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Prepro-mGnRH precursor (Gonadoliberin) (Gonadotropin-releasing
DE   hormone) (LH-RH) (Luliberin).
OS   Anguilla japonica (Japanese eel).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Anguilliformes; Anguillidae;
OC   Anguilla.
OX   NCBI_TaxID=7937;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RA   Okubo K., Suetake H., Aida K.;
RT   "Expression of two gonadotropin-releasing hormone (GnRH) precursor
RT   genes in various tissues of the Japanese eel and evolution of GnRH.";
RL   Zool. Sci. 16:471-478(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Okubo K., Suetake H., Aida K.;
RT   "A splicing variant for the prepro-mammalian gonadotropin-releasing
RT   hormone (prepro-mGnRH) mRNA is present in the brain and various
RT   peripheral tissues of the Japanese eel.";
RL   Zool. Sci. 16:645-651(1999).
CC   -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY
CC       SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR   EMBL; AB026989; BAA82608.1; -.
DR   EMBL; AB026991; BAA83597.1; -.
DR   GO; GO:0005576; C:extracellular; IEA.
DR   GO; GO:0005183; F:luteinizing hormone-releasing factor activity; IEA.
DR   GO; GO:0007275; P:development; IEA.
DR   InterPro; IPR002012; GnRH.
DR   InterPro; IPR004079; GonadoliberinI.
DR   Pfam; PF00446; GnRH; 1.
DR   PRINTS; PR01541; GONADOLIBRNI.
DR   PROSITE; PS00473; GNRH; 1.
KW   Amidation; Hormone; Signal.
FT   SIGNAL        1     22       POTENTIAL.
FT   CHAIN        23     32       MGNRH.
FT   CHAIN        33     91       GNRH ASSOCIATED PEPTIDE.
SQ   SEQUENCE   91 AA;  9893 MW;  BA15C9DC08434A7B CRC64;

  Query Match             29.0%;  Score 58;  DB 13;  Length 91;
  Best Local Similarity   100.0%;  Pred. No. 1.6;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 HWSYGLRPG 10
              |||||||||
Db         24 HWSYGLRPG 32


RESULT 14
Q9W9A5
ID   Q9W9A5      PRELIMINARY;      PRT;   234 AA.
AC   Q9W9A5;
```

```
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Viral coat protein 3.
OS   Simian virus 40 (SV40).
OC   Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
OX   NCBI_TaxID=10633;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=VA45-54-1, and VA45-54-2;
RX   MEDLINE=97368433; PubMed=9225047;
RA   Lednicky J.A., Butel J.S.;
RT   "Tissue culture adaptation of natural isolates of simian virus 40:
RT   changes occur in viral regulatory region but not in carboxy-terminal
RT   domain of large T-antigen.";
RL   J. Gen. Virol. 78:0-0(0).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=VA45-54-1, and VA45-54-2;
RA   Lednicky J.A., Stewart A.R., Butel J.S.;
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=VA45-54-2;
RX   MEDLINE=98244346; PubMed=9584955;
RA   Stewart A.R., Lednicky J.A., Butel J.S.;
RT   "Sequence analyses of human tumor-associated SV40 DNAs and SV40 viral
RT   isolates from monkeys and humans.";
RL   J. Neurovirol. 4:182-193(1998).
DR   EMBL; AF156107; AAD38991.1; -.
DR   EMBL; AF156105; AAD38844.1; -.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   InterPro; IPR001070; Polyoma_coat2.
DR   Pfam; PF00761; Polyoma_coat2; 1.
KW   Coat protein.
SQ   SEQUENCE   234 AA;  26962 MW;  24904AD332616377 CRC64;

  Query Match             28.7%;  Score 57.5;  DB 12;  Length 234;
  Best Local Similarity   30.0%;  Pred. No. 5.4;
  Matches   12;  Conservative    8;  Mismatches    7;  Indels   13;  Gaps    2;

Qy          2 HWSYGLRPGSSGPSLFNNFTVSFWL----RVPKVSASHLE 37
              ||         ||:|||  : :||      :|::::  ||
Db         33 HW---------GPTLFNAISQAFWRVIQNDIPRLTSQELE 63


RESULT 15
O92837
ID   O92837      PRELIMINARY;      PRT;   234 AA.
AC   O92837;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   VP3 (VP3 minor structural protein) (Minor structural protein VP3)
DE   (Viral coat protein 3) (Minor viral coat protein 3).
GN   VP3.
OS   Simian virus 40 (SV40).
OC   Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
OX   NCBI_TaxID=10633;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K661;
RX   MEDLINE=98216763; PubMed=9557685;
RA   Lednicky J.A., Arrington A.S., Stewart A.R., Cai X.M., Wong C.,
RA   Jafar S., Murphey-Corb M., Butel J.S.;
RT   "Natural isolates of simian virus 40 from immunocompromised monkeys
RT   display extensive genetic heterogeneity: new implications for
RT   polyomavirus disease.";
RL   J. Virol. 72:3980-3990(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K661;
RA   Lednicky J.A., Arrington A.S., Stewart A.R., Dai X.M., Butel J.S.;
RL   Submitted (DEC-1997) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=B-1, and B-2;
RX   MEDLINE=97368433; PubMed=9225047;
RA   Lednicky J.A., Butel J.S.;
RT   "Tissue culture adaptation of natural isolates of simian virus 40:
RT   changes occur in viral regulatory region but not in carboxy-terminal
RT   domain of large T-antigen.";
RL   J. Gen. Virol. 78:0-0(0).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=B-1, CPC/MEN, and B-2;
RX   MEDLINE=98244346; PubMed=9584955;
RA   Stewart A.R., Lednicky J.A., Butel J.S.;
RT   "Sequence analyses of human tumor-associated SV40 DNAs and SV40 viral
RT   isolates from monkeys and humans.";
RL   J. Neurovirol. 4:182-193(1998).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=B-1, and B-2;
RA   Lednicky J.A., Stewart A.R., Butel J.S.;
RL   Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CPC/MEN;
RX   MEDLINE=96010244; PubMed=7571441;
RA   Lednicky J.A., Garcea R.L., Bergsagel D.J., Butel J.S.;
RT   "Natural simian virus 40 strains are present in human choroid plexus
RT   and ependymoma tumors.";
RL   Virology 212:710-717(1995).
RN   [7]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CPC/MEN;
RA   Lednicky J.A., Stewart A.R., Butel J.S.;
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
RN   [8]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MC-028846B;
RA   Rizzo P., Carbone M.;
RT   "SV40 from 1955 commercial parenteral (Salk) poliovaccine.";
RL   Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
RN   [9]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H328;
RA   Lednicky J.A., Butel J.S., Lewis A.M.;
RT   "Complete DNA sequence of SV40 strain H328-1, a regulatory region
RT   variant of SV40 strain H328 with an archetypal regulatory region.";
RL   Submitted (OCT-2000) to the EMBL/GenBank/DDBJ databases.
RN   [10]
RP   SEQUENCE FROM N.A.
RC   STRAIN=776;
RA   Lednicky J.A., Butel J.S., Lewis A.M.;
RT   "DNA sequence of SV40 reference strain 776.";
RL   Submitted (OCT-2000) to the EMBL/GenBank/DDBJ databases.
RN   [11]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Rh911;
RA   Lednicky J.A., Butel J.S., Lewis A.M.;
RT   "DNA sequence of SV40 strain Rh911, regulatory region variant 1.";
RL   Submitted (OCT-2000) to the EMBL/GenBank/DDBJ databases.
RN   [12]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GM00637H;
RA   Lednicky J.A., Butel J.S., Lewis A.M.;
RT   "Complete DNA sequence of SV40-GM00637H, defective variant 12.";
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
RN   [13]
RP   SEQUENCE FROM N.A.
RC   STRAIN=777;
RA   Lednicky J.A., Butel J.S., Lewis A.M.;
RT   "Complete DNA sequence of SV40 strain 777, regulatory region variant
```

```
RT   5.";
RL   Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
RN   [14]
RP   SEQUENCE FROM N.A.
RC   STRAIN=777;
RA   Lednicky J.A., Butel J.S., Lewis A.M.;
RT   "Complete DNA sequence of SV40 strain 777, regulatory region variant
RT   1.";
RL   Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
RN   [15]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GM00637H;
RA   Lednicky J.A., Butel J.S., Lewis A.M.;
RT   "Complete DNA Sequence of SV40-GM00637H Variant 1.";
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
RN   [16]
RP   SEQUENCE FROM N.A.
RC   STRAIN=N128-1;
RA   Lednicky J.A., Butel J.S., Lewis A.M.;
RT   "Whole genomic sequence of SV-40 isolate N128-1.";
RL   Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.
RN   [17]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PML-1EK;
RA   Lednicky J.A., Stewart A.R., Butel J.S., Lewis A.M.;
RT   "Full-length genomic sequence of SV40-PML-1 EK(alpha regulatory
RT   region).";
RL   Submitted (APR-2003) to the EMBL/GenBank/DDBJ databases.
RN   [18]
RP   SEQUENCE FROM N.A.
RC   STRAIN=777*;
RA   Lednicky J.A., Butel J.S., Lewis A.M.;
RT   "Complete genomic sequence of SV40 strain 777*.";
RL   Submitted (APR-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF038616; AAC59344.1; -.
DR   EMBL; AF155359; AAD43803.1; -.
DR   EMBL; AF156108; AAD38999.1; -.
DR   EMBL; AF155358; AAD43797.1; -.
DR   EMBL; AF180737; AAF28268.1; -.
DR   EMBL; AF316141; AAG39212.1; -.
DR   EMBL; AF316139; AAG39200.1; -.
DR   EMBL; AF316140; AAG39206.1; -.
DR   EMBL; AF345345; AAK29052.1; -.
DR   EMBL; AF332699; AAK01716.1; -.
DR   EMBL; AF332562; AAK19523.1; -.
DR   EMBL; AF345344; AAK29046.1; -.
DR   EMBL; AY120890; AAM77805.1; -.
DR   EMBL; AY271816; AAP30058.1; -.
DR   EMBL; AY271817; AAP30064.1; -.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   InterPro; IPR001070; Polyoma_coat2.
DR   Pfam; PF00761; Polyoma_coat2; 1.
KW   Coat protein.
SQ   SEQUENCE   234 AA;  26963 MW;  FFF86A591AC3E957 CRC64;

  Query Match             28.7%;  Score 57.5;  DB 12;  Length 234;
  Best Local Similarity   30.0%;  Pred. No. 5.4;
  Matches   12;  Conservative    8;  Mismatches    7;  Indels   13;  Gaps    2;

Qy          2 HWSYGLRPGSSGPSLFNNFTVSFWL----RVPKVSASHLE 37
              ||          ||:|||  : :||      :|::::  ||
Db         33 HW---------GPTLFNAISQAFWRVIQNDIPRLTSQELE 63


Search completed: March 10, 2004, 09:25:36
Job time : 37.5681 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:54 ; Search time 13.1634 Seconds
(without alignments)
133.345 Million cell updates/sec

Title: US-09-848-834A-13
Perfect score: 174
Sequence: 1 XHWSYGLRPGSSGPSLKLLSEIKGVIVHRLEGVE 34

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	106	60.9	40	2	US-08-460-502-10	Sequence 10, Appl
2	106	60.9	40	4	US-09-148-711A-10	Sequence 10, Appl
3	101	58.0	40	2	US-08-460-502-11	Sequence 11, Appl
4	96	55.2	40	4	US-09-148-711A-11	Sequence 11, Appl
5	73.5	42.2	34	5	PCT-US95-13841-28	Sequence 28, Appl
6	73.5	42.2	43	5	PCT-US95-13841-26	Sequence 26, Appl
7	73	42.0	42	1	US-08-446-692-20	Sequence 20, Appl
8	73	42.0	42	2	US-08-488-351A-20	Sequence 20, Appl
9	72	41.4	15	3	US-09-100-414B-1	Sequence 1, Appli
10	72	41.4	15	3	US-09-100-409A-38	Sequence 38, Appl
11	72	41.4	15	3	US-09-303-323-1	Sequence 1, Appli
12	72	41.4	15	4	US-09-770-014-1	Sequence 1, Appli
13	72	41.4	15	5	PCT-US95-13841-12	Sequence 12, Appl
14	72	41.4	27	1	US-08-446-692-19	Sequence 19, Appl
15	72	41.4	27	2	US-08-488-351A-19	Sequence 19, Appl
16	72	41.4	27	3	US-09-100-414B-36	Sequence 36, Appl
17	72	41.4	27	3	US-09-303-323-36	Sequence 36, Appl
18	72	41.4	27	4	US-09-770-014-36	Sequence 36, Appl
19	72	41.4	34	5	PCT-US95-13841-27	Sequence 27, Appl
20	72	41.4	35	1	US-08-446-692-55	Sequence 55, Appl
21	72	41.4	35	1	US-08-446-692-61	Sequence 61, Appl
22	72	41.4	35	2	US-08-488-351A-55	Sequence 55, Appl
23	72	41.4	35	2	US-08-488-351A-61	Sequence 61, Appl
24	72	41.4	36	3	US-09-082-279B-505	Sequence 505, App
25	72	41.4	36	4	US-09-315-304B-505	Sequence 505, App
26	72	41.4	36	4	US-09-834-784-505	Sequence 505, App
27	72	41.4	36	4	US-09-515-965A-505	Sequence 505, App
28	72	41.4	36	4	US-09-350-641C-505	Sequence 505, App
29	72	41.4	42	5	PCT-US95-13841-25	Sequence 25, Appl
30	72	41.4	45	1	US-08-446-692-33	Sequence 33, Appl
31	72	41.4	45	2	US-08-488-351A-33	Sequence 33, Appl
32	72	41.4	438	3	US-08-486-099-105	Sequence 105, App
33	72	41.4	438	3	US-08-360-107A-115	Sequence 115, App
34	72	41.4	438	3	US-08-484-223B-105	Sequence 105, App
35	72	41.4	438	3	US-08-919-597-105	Sequence 105, App
36	72	41.4	438	3	US-08-475-668A-105	Sequence 105, App
37	72	41.4	438	3	US-08-485-551A-105	Sequence 105, App
38	72	41.4	438	3	US-08-471-913A-105	Sequence 105, App
39	72	41.4	438	3	US-08-485-264A-105	Sequence 105, App
40	72	41.4	438	4	US-08-474-349A-105	Sequence 105, App
41	72	41.4	438	4	US-08-470-896-105	Sequence 105, App
42	72	41.4	438	4	US-08-485-546A-105	Sequence 105, App
43	72	41.4	550	1	US-08-279-700-16	Sequence 16, Appl
44	72	41.4	550	1	US-08-279-700-18	Sequence 18, Appl
45	72	41.4	550	1	US-08-279-700-20	Sequence 20, Appl

ALIGNMENTS

```
RESULT 1
US-08-460-502-10
; Sequence 10, Application US/08460502
; Patent No. 5843464
;  GENERAL INFORMATION:
;    APPLICANT:  Bakaletz, Lauren O.
;    APPLICANT:  Kaumaya, Parvin T.
;    TITLE OF INVENTION:  Synthetic Chimeric Fimbrin Peptides
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Calfee, Halter and Griswold
;      STREET:  800 Superior Avenue
;      CITY:  Cleveland
;      STATE:  Ohio
;      COUNTRY:  U.S.A.
;      ZIP:  44114-2688
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/460,502
;      FILING DATE:
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Golrick, Mary E.
;      REGISTRATION NUMBER:  34,829
;      REFERENCE/DOCKET NUMBER:  22727/00120
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (216) 622-8458
;      TELEFAX:  (216) 241-0816
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  40 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-460-502-10

  Query Match             60.9%;  Score 106;  DB 2;  Length 40;
  Best Local Similarity   79.3%;  Pred. No. 9.7e-09;
  Matches   23;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          6 GLRPGSSGPSLKLLSEIKGVIVHRLEGVE 34
              | |    |||||||| ||||||||||||
Db         12 GTRDHKKGPSLKLLSLIKGVIVHRLEGVE 40
```

```
RESULT 2
US-09-148-711A-10
; Sequence 10, Application US/09148711A
; Patent No. 6436405
; GENERAL INFORMATION:
;  APPLICANT:  The Ohio State University
;  TITLE OF INVENTION:   Synthetic Chimeric Fimbrin Peptides
;  FILE REFERENCE:   18525-04010
;  CURRENT APPLICATION NUMBER: US/09/148,711A
;  CURRENT FILING DATE:  1998-09-04
;  PRIOR APPLICATION NUMBER:   08/460,502
;  PRIOR FILING DATE:   1995-06-02
;  NUMBER OF SEQ ID NOS:   12
;  SOFTWARE:   PatentIn version 3.0
; SEQ ID NO 10
;   LENGTH: 40
;   TYPE: PRT
;   ORGANISM: synthetic construct
US-09-148-711A-10

  Query Match             60.9%;  Score 106;  DB 4;  Length 40;
  Best Local Similarity   79.3%;  Pred. No. 9.7e-09;
  Matches   23;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          6 GLRPGSSGPSLKLLSEIKGVIVHRLEGVE 34
              | |    |||||||| ||||||||||||
Db         12 GTRDHKKGPSLKLLSLIKGVIVHRLEGVE 40


RESULT 3
US-08-460-502-11
; Sequence 11, Application US/08460502
; Patent No. 5843464
;  GENERAL INFORMATION:
;    APPLICANT:  Bakaletz, Lauren O.
;    APPLICANT:  Kaumaya, Parvin T.
;    TITLE OF INVENTION:  Synthetic Chimeric Fimbrin Peptides
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Calfee, Halter and Griswold
;      STREET:  800 Superior Avenue
;      CITY:  Cleveland
;      STATE:  Ohio
;      COUNTRY:  U.S.A.
;      ZIP:  44114-2688
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/460,502
;      FILING DATE:
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Golrick, Mary E.
;      REGISTRATION NUMBER:  34,829
;      REFERENCE/DOCKET NUMBER:  22727/00120
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (216) 622-8458
;      TELEFAX:  (216) 241-0816
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  40 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-460-502-11

  Query Match             58.0%;  Score 101;  DB 2;  Length 40;
  Best Local Similarity   87.5%;  Pred. No. 5.3e-08;
  Matches   21;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         11 SSGPSLKLLSEIKGVIVHRLEGVE 34
              ::|||||||| ||||||||||||
Db         17 NTGPSLKLLSLIKGVIVHRLEGVE 40


RESULT 4
US-09-148-711A-11
; Sequence 11, Application US/09148711A
; Patent No. 6436405
; GENERAL INFORMATION:
;  APPLICANT:  The Ohio State University
;  TITLE OF INVENTION:   Synthetic Chimeric Fimbrin Peptides
;  FILE REFERENCE:   18525-04010
;  CURRENT APPLICATION NUMBER: US/09/148,711A
;  CURRENT FILING DATE:  1998-09-04
;  PRIOR APPLICATION NUMBER:   08/460,502
;  PRIOR FILING DATE:   1995-06-02
;  NUMBER OF SEQ ID NOS:   12
;  SOFTWARE:   PatentIn version 3.0
; SEQ ID NO 11
;   LENGTH: 40
;   TYPE: PRT
;   ORGANISM: SYNTHETIC CONSTRUCT
US-09-148-711A-11

  Query Match             55.2%;  Score 96;  DB 4;  Length 40;
  Best Local Similarity   87.0%;  Pred. No. 2.9e-07;
  Matches   20;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         11 SSGPSLKLLSEIKGVIVHRLEGV 33
              ::|||||||| ||||||||||||
Db         17 NTGPSLKLLSLIKGVIVHRLEGV 39


RESULT 5
PCT-US95-13841-28
; Sequence 28, Application PC/TUS9513841
;  GENERAL INFORMATION:
;    APPLICANT:  United Biomedical Inc; Walfield, Alan M.;
;    APPLICANT:  Wang, Chang Yi
;    TITLE OF INVENTION:  Synthetic IgE Membrane Anchor
;    TITLE OF INVENTION:  Peptide Immunogens for the Treatment of Allergy
;    NUMBER OF SEQUENCES:  32
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Maria C.H. Lin
;      STREET:  345 Park Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10154
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/13841
;      FILING DATE:  25-OCT-1995
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/328,519
;      FILING DATE:  25-OCT-1994
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lin, Maria C.H.
;      REGISTRATION NUMBER:  29,323
;      REFERENCE/DOCKET NUMBER:  1151-4117
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-758-4800
;      TELEFAX:  212-751-6849
;      TELEX:  421792
;  INFORMATION FOR SEQ ID NO:  28:
```

```
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  34 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
PCT-US95-13841-28

  Query Match             42.2%;  Score 73.5;  DB 5;  Length 34;
  Best Local Similarity   64.3%;  Pred. No. 0.00052;
  Matches   18;  Conservative    1;  Mismatches    4;  Indels    5;  Gaps    1;

Qy          6 GLRPGSSGPSLKLLSEIKGVIVHRLEGV 33
              |  | : |     ||||||||||||||
Db         12 GEAPWTGG-----LSEIKGVIVHRLEGV 34


RESULT 6
PCT-US95-13841-26
; Sequence 26, Application PC/TUS9513841
;  GENERAL INFORMATION:
;    APPLICANT:  United Biomedical Inc; Walfield, Alan M.;
;    APPLICANT:  Wang, Chang Yi
;    TITLE OF INVENTION:  Synthetic IgE Membrane Anchor
;    TITLE OF INVENTION:  Peptide Immunogens for the Treatment of Allergy
;    NUMBER OF SEQUENCES:  32
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Maria C.H. Lin
;      STREET:  345 Park Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10154
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/13841
;      FILING DATE:  25-OCT-1995
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/328,519
;      FILING DATE:  25-OCT-1994
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lin, Maria C.H.
;      REGISTRATION NUMBER:  29,323
;      REFERENCE/DOCKET NUMBER:  1151-4117
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-758-4800
;      TELEFAX:  212-751-6849
;      TELEX:  421792
;  INFORMATION FOR SEQ ID NO:  26:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  43 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
PCT-US95-13841-26

  Query Match             42.2%;  Score 73.5;  DB 5;  Length 43;
  Best Local Similarity   64.3%;  Pred. No. 0.00069;
  Matches   18;  Conservative    1;  Mismatches    4;  Indels    5;  Gaps    1;

Qy          6 GLRPGSSGPSLKLLSEIKGVIVHRLEGV 33
              | : |  |     ||||||||||||||
Db         21 GQQQGLGG-----LSEIKGVIVHRLEGV 43


RESULT 7
US-08-446-692-20
; Sequence 20, Application US/08446692
; Patent No. 5759551
;  GENERAL INFORMATION:
;    APPLICANT:  Ladd, Anna
;    APPLICANT:  Wang, Chang Yi
;    APPLICANT:  Zamb, Timothy
;    TITLE OF INVENTION:  Immunogenic LHRH peptide constructs
;    TITLE OF INVENTION:  and synthetic universal immune stimulators for vaccines
;    NUMBER OF SEQUENCES:  114
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Maria C.H. Lin
;      STREET:  345 Park Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  US
;      ZIP:  10154-0053
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/446,692
;      FILING DATE:  7-JUN-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maria C.H. Lin
;      REGISTRATION NUMBER:  29,323
;      REFERENCE/DOCKET NUMBER:  1151-4146 US2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212)415-8745
;      TELEFAX:  (516)751-6849
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  42 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-446-692-20

  Query Match             42.0%;  Score 73;  DB 1;  Length 42;
  Best Local Similarity   93.8%;  Pred. No. 0.00079;
  Matches   15;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         18 LLSEIKGVIVHRLEGV 33
              :|||||||||||||||
Db         15 VLSEIKGVIVHRLEGV 30


RESULT 8
US-08-488-351A-20
; Sequence 20, Application US/08488351A
; Patent No. 5843446
;  GENERAL INFORMATION:
;    APPLICANT:  Ladd, Anna
;    APPLICANT:  Wang, Chang Yi
;    APPLICANT:  Zamb, Timothy
;    TITLE OF INVENTION:  Immunogenic LHRH peptide constructs
;    TITLE OF INVENTION:  and synthetic universal immune stimulators for vaccines
;    NUMBER OF SEQUENCES:  114
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Maria C.H. Lin
;      STREET:  345 Park Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  US
;      ZIP:  10154-0053
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/488,351A
```

```
;      FILING DATE:  7-JUN-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/446,692
;      FILING DATE:  7-JUN-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/229,275
;      FILING DATE:  14-APR-1994
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/057,166
;      FILING DATE:  27-APR-1992
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maria C.H. Lin
;      REGISTRATION NUMBER:  29,323
;      REFERENCE/DOCKET NUMBER:  1151-4146 US2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212)415-8745
;      TELEFAX:  (516)751-6849
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  42 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-488-351A-20

  Query Match             42.0%;  Score 73;  DB 2;  Length 42;
  Best Local Similarity   93.8%;  Pred. No. 0.00079;
  Matches   15;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         18 LLSEIKGVIVHRLEGV 33
              :||||||||||||||||
Db         15 VLSEIKGVIVHRLEGV 30


RESULT 9
US-09-100-414B-1
; Sequence 1, Application US/09100414B
; Patent No. 6025468
;  GENERAL INFORMATION:
;    APPLICANT:  Wang, Chang Yi
;    TITLE OF INVENTION:  NOVEL LHRH PEPTIDE
;    TITLE OF INVENTION:  IMMUNOGENS
;    NUMBER OF SEQUENCES:  106
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morgan & Finnegan, L.L.P.
;      STREET:  345 Park Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10154-0054
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC Windows
;      SOFTWARE:  Word 97
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/100,414B
;      FILING DATE:  20-JUNE-1998
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maria H. Lin
;      REGISTRATION NUMBER:  29,323
;      REFERENCE/DOCKET NUMBER:  1151-4157
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-758-4800
;      TELEFAX:  212-751-6849
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  15 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-100-414B-1

  Query Match             41.4%;  Score 72;  DB 3;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 0.00032;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         19 LSEIKGVIVHRLEGV 33
              |||||||||||||||
Db          1 LSEIKGVIVHRLEGV 15


RESULT 10
US-09-100-409A-38
; Sequence 38, Application US/09100409A
; Patent No. 6090388
;  GENERAL INFORMATION:
;    APPLICANT:  Wang, Chang Yi
;    TITLE OF INVENTION:  PEPTIDE COMPOSITION FOR
;    TITLE OF INVENTION:  PREVENTION AND TREATMENT OF HIV INFECTION AND
;    TITLE OF INVENTION:  IMMUNE DISORDERS
;    NUMBER OF SEQUENCES:  64
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORGAN & FINNEGAN
;      STREET:  345 Park Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10154-0054
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version
;      SOFTWARE:  #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/100,409A
;      FILING DATE:
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:
;      REGISTRATION NUMBER:
;      REFERENCE/DOCKET NUMBER:  1151-4154
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-758-4800
;      TELEFAX:  212-751-6849
;  INFORMATION FOR SEQ ID NO:  38:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  15 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-100-409A-38

  Query Match             41.4%;  Score 72;  DB 3;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 0.00032;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         19 LSEIKGVIVHRLEGV 33
              |||||||||||||||
Db          1 LSEIKGVIVHRLEGV 15


RESULT 11
US-09-303-323-1
; Sequence 1, Application US/09303323
; Patent No. 6228987
;  GENERAL INFORMATION:
;    APPLICANT:  Wang, Chang Yi
;    TITLE OF INVENTION:  NOVEL LHRH PEPTIDE
;    TITLE OF INVENTION:  IMMUNOGENS
```

```
;    NUMBER OF SEQUENCES:  106
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morgan & Finnegan, L.L.P.
;      STREET:  345 Park Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10154-0054
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC Windows
;      SOFTWARE:  Word 97
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/303,323
;      FILING DATE:  30-APR-1999
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/100,414
;      FILING DATE:  20-JUNE-1998
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maria H. Lin
;      REGISTRATION NUMBER:  29,323
;      REFERENCE/DOCKET NUMBER:  1151-4157
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-758-4800
;      TELEFAX:  212-751-6849
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  15 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-303-323-1

  Query Match             41.4%;  Score 72;  DB 3;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 0.00032;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         19 LSEIKGVIVHRLEGV 33
              |||||||||||||||
Db          1 LSEIKGVIVHRLEGV 15


RESULT 12
US-09-770-014-1
; Sequence 1, Application US/09770014
; Patent No. 6559282
;  GENERAL INFORMATION:
;    APPLICANT:  Wang, Chang Yi
;    TITLE OF INVENTION:  NOVEL LHRH PEPTIDE
;    TITLE OF INVENTION:  IMMUNOGENS
;    NUMBER OF SEQUENCES:  106
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morgan & Finnegan, L.L.P.
;      STREET:  345 Park Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10154-0054
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC Windows
;      SOFTWARE:  Word 97
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/770,014
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/100,414
;      FILING DATE:  20-JUNE-1998
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maria H. Lin
;      REGISTRATION NUMBER:  29,323
;      REFERENCE/DOCKET NUMBER:  1151-4157
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-758-4800
;      TELEFAX:  212-751-6849
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  15 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-770-014-1

  Query Match             41.4%;  Score 72;  DB 4;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 0.00032;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         19 LSEIKGVIVHRLEGV 33
              |||||||||||||||
Db          1 LSEIKGVIVHRLEGV 15


RESULT 13
PCT-US95-13841-12
; Sequence 12, Application PC/TUS9513841
;  GENERAL INFORMATION:
;    APPLICANT:  United Biomedical Inc; Walfield, Alan M.;
;    APPLICANT:  Wang, Chang Yi
;    TITLE OF INVENTION:  Synthetic IgE Membrane Anchor
;    TITLE OF INVENTION:  Peptide Immunogens for the Treatment of Allergy
;    NUMBER OF SEQUENCES:  32
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Maria C.H. Lin
;      STREET:  345 Park Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10154
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/13841
;      FILING DATE:  25-OCT-1995
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/328,519
;      FILING DATE:  25-OCT-1994
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lin, Maria C.H.
;      REGISTRATION NUMBER:  29,323
;      REFERENCE/DOCKET NUMBER:  1151-4117
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-758-4800
;      TELEFAX:  212-751-6849
;      TELEX:  421792
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  15 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
PCT-US95-13841-12

  Query Match             41.4%;  Score 72;  DB 5;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 0.00032;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         19 LSEIKGVIVHRLEGV 33
```

```
                  ||||||||||||||||
Db          1 LSEIKGVIVHRLEGV 15


RESULT 14
US-08-446-692-19
; Sequence 19, Application US/08446692
; Patent No. 5759551
;  GENERAL INFORMATION:
;    APPLICANT:  Ladd, Anna
;    APPLICANT:  Wang, Chang Yi
;    APPLICANT:  Zamb, Timothy
;    TITLE OF INVENTION:  Immunogenic LHRH peptide constructs
;    TITLE OF INVENTION:  and synthetic universal immune stimulators for vaccines
;    NUMBER OF SEQUENCES:  114
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Maria C.H. Lin
;      STREET:  345 Park Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  US
;      ZIP:  10154-0053
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/446,692
;      FILING DATE:  7-JUN-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maria C.H. Lin
;      REGISTRATION NUMBER:  29,323
;      REFERENCE/DOCKET NUMBER:  1151-4146 US2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212)415-8745
;      TELEFAX:  (516)751-6849
;  INFORMATION FOR SEQ ID NO:  19:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  27 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-446-692-19

  Query Match             41.4%;  Score 72;  DB 1;  Length 27;
  Best Local Similarity   100.0%;  Pred. No. 0.00065;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         19 LSEIKGVIVHRLEGV 33
                ||||||||||||||
Db          1 LSEIKGVIVHRLEGV 15


RESULT 15
US-08-488-351A-19
; Sequence 19, Application US/08488351A
; Patent No. 5843446
;  GENERAL INFORMATION:
;    APPLICANT:  Ladd, Anna
;    APPLICANT:  Wang, Chang Yi
;    APPLICANT:  Zamb, Timothy
;    TITLE OF INVENTION:  Immunogenic LHRH peptide constructs
;    TITLE OF INVENTION:  and synthetic universal immune stimulators for vaccines
;    NUMBER OF SEQUENCES:  114
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Maria C.H. Lin
;      STREET:  345 Park Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  US
;      ZIP:  10154-0053
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/488,351A
;      FILING DATE:  7-JUN-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/446,692
;      FILING DATE:  7-JUN-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/229,275
;      FILING DATE:  14-APR-1994
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/057,166
;      FILING DATE:  27-APR-1992
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maria C.H. Lin
;      REGISTRATION NUMBER:  29,323
;      REFERENCE/DOCKET NUMBER:  1151-4146 US2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212)415-8745
;      TELEFAX:  (516)751-6849
;  INFORMATION FOR SEQ ID NO:  19:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  27 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-488-351A-19

  Query Match             41.4%;  Score 72;  DB 2;  Length 27;
  Best Local Similarity   100.0%;  Pred. No. 0.00065;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         19 LSEIKGVIVHRLEGV 33
                ||||||||||||||
Db          1 LSEIKGVIVHRLEGV 15


Search completed: March 10, 2004, 09:28:55
Job time : 14.1634 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 10, 2004, 09:16:59 ; Search time 26.7237 Seconds
(without alignments)
268.645 Million cell updates/sec

Title: US-09-848-834A-13
Perfect score: 174
Sequence: 1 XHWSYGLRPGSSGPSLKLLSEIKGVIVHRLEGVE 34

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	173	99.4	34	9	US-09-848-834A-13	Sequence 13, Appl
2	173	99.4	47	9	US-09-848-834A-17	Sequence 17, Appl
3	106	60.9	40	14	US-10-223-711-10	Sequence 10, Appl
4	102	58.6	75	10	US-09-847-102A-31	Sequence 31, Appl
5	102	58.6	75	14	US-10-285-976-229	Sequence 229, App
6	96	55.2	40	14	US-10-223-711-11	Sequence 11, Appl
7	89	51.1	31	9	US-09-848-834A-15	Sequence 15, Appl
8	89	51.1	46	9	US-09-848-834A-19	Sequence 19, Appl
9	87	50.0	36	9	US-09-848-834A-16	Sequence 16, Appl
10	87	50.0	37	9	US-09-848-834A-14	Sequence 14, Appl
11	87	50.0	50	9	US-09-848-834A-18	Sequence 18, Appl
12	87	50.0	51	9	US-09-848-834A-20	Sequence 20, Appl
13	86	49.4	31	9	US-09-848-834A-9	Sequence 9, Appli
14	80	46.0	75	10	US-09-847-102A-33	Sequence 33, Appl
15	80	46.0	75	14	US-10-285-976-231	Sequence 231, App
16	77	44.3	16	9	US-09-848-834A-8	Sequence 8, Appli
17	72	41.4	15	10	US-09-747-802-16	Sequence 16, Appl
18	72	41.4	15	10	US-09-747-802-30	Sequence 30, Appl
19	72	41.4	15	10	US-09-865-294-8	Sequence 8, Appli
20	72	41.4	15	10	US-09-865-294-22	Sequence 22, Appl
21	72	41.4	15	14	US-10-261-446-20	Sequence 20, Appl
22	72	41.4	15	15	US-10-411-544-10	Sequence 10, Appl
23	72	41.4	25	15	US-10-411-544-32	Sequence 32, Appl
24	72	41.4	36	14	US-10-351-641-505	Sequence 505, App
25	72	41.4	45	14	US-10-076-674-9	Sequence 9, Appli
26	72	41.4	45	15	US-10-355-161A-9	Sequence 9, Appli
27	72	41.4	550	9	US-09-873-233A-18	Sequence 18, Appl
28	72	41.4	550	9	US-09-873-233A-20	Sequence 20, Appl
29	71	40.8	17	10	US-09-305-924-4	Sequence 4, Appli
30	69	39.7	15	10	US-09-747-802-37	Sequence 37, Appl
31	69	39.7	15	10	US-09-865-294-29	Sequence 29, Appl
32	69	39.7	16	14	US-10-223-711-3	Sequence 3, Appli
33	69	39.7	19	10	US-09-747-802-48	Sequence 48, Appl
34	69	39.7	19	10	US-09-865-294-40	Sequence 40, Appl
35	66	37.9	19	10	US-09-747-802-51	Sequence 51, Appl
36	66	37.9	19	10	US-09-865-294-43	Sequence 43, Appl
37	66	37.9	662	10	US-09-951-061A-141	Sequence 141, App
38	64	36.8	15	10	US-09-747-802-35	Sequence 35, Appl
39	64	36.8	15	10	US-09-747-802-38	Sequence 38, Appl
40	64	36.8	15	10	US-09-747-802-40	Sequence 40, Appl
41	64	36.8	15	10	US-09-747-802-42	Sequence 42, Appl
42	64	36.8	15	10	US-09-865-294-27	Sequence 27, Appl
43	64	36.8	15	10	US-09-865-294-30	Sequence 30, Appl
44	64	36.8	15	10	US-09-865-294-32	Sequence 32, Appl
45	64	36.8	15	10	US-09-865-294-34	Sequence 34, Appl

ALIGNMENTS

RESULT 1
US-09-848-834A-13
; Sequence 13, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Aphton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 13
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the
; OTHER INFORMATION: RH hormone linked by a spacer to amino acid sequence 288-302 o
; OTHER INFORMATION: he Measles virus fusion protein,
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(10)
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
; NAME/KEY: PEPTIDE
; LOCATION: (11)..(18)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (19)..(34)
; OTHER INFORMATION: Amino acid sequence 288-302 of the Measles
; OTHER INFORMATION: virus fusion protein, F
; NAME/KEY: MOD_RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
US-09-848-834A-13

Query Match 99.4%; Score 173; DB 9; Length 34;

```
  Best Local Similarity   100.0%;  Pred. No. 3.3e-17;
  Matches   33;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 HWSYGLRPGSSGPSLKLLSEIKGVIVHRLEGVE 34
              |||||||||||||||||||||||||||||||||
Db          2 HWSYGLRPGSSGPSLKLLSEIKGVIVHRLEGVE 34


RESULT 2
US-09-848-834A-17
; Sequence 17, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
;  APPLICANT: Aphton Corporation
;  TITLE OF INVENTION: Chimeric Peptide Immunogens
;  FILE REFERENCE: 1102865-0047
;  CURRENT APPLICATION NUMBER: US/09/848,834A
;  CURRENT FILING DATE:  2001-05-04
;  PRIOR APPLICATION NUMBER: 60/202,328
;  PRIOR FILING DATE: 2000-05-05
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
;   LENGTH: 47
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the Gn
;   OTHER INFORMATION: RH hormone linked by a spacer to amino acid sequence 288-302 of
;   OTHER INFORMATION: the Measles virus protein F linked by a spacer to amino acid seq
;   OTHER INFORMATION: uence 2-10 of the GnRH hormone
;   NAME/KEY: MOD_RES
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
;   NAME/KEY: MOD_RES
;   LOCATION: (47)..(47)
;   OTHER INFORMATION: Amidated-glycine or glycinamide
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)..(10)
;   OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
;   NAME/KEY: PEPTIDE
;   LOCATION: (11)..(18)
;   OTHER INFORMATION: Spacer peptide
;   NAME/KEY: PEPTIDE
;   LOCATION: (19)..(34)
;   OTHER INFORMATION: Amino acid sequence 288-302 of the Measles virus fusion protein,
;   NAME/KEY: PEPTIDE
;   LOCATION: (35)..(38)
;   OTHER INFORMATION: Spacer peptide
;   NAME/KEY: PEPTIDE
;   LOCATION: (39)..(47)
;   OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone
US-09-848-834A-17

  Query Match             99.4%;  Score 173;  DB 9;  Length 47;
  Best Local Similarity   100.0%;  Pred. No. 4.7e-17;
  Matches   33;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 HWSYGLRPGSSGPSLKLLSEIKGVIVHRLEGVE 34
              |||||||||||||||||||||||||||||||||
Db          2 HWSYGLRPGSSGPSLKLLSEIKGVIVHRLEGVE 34


RESULT 3
US-10-223-711-10
; Sequence 10, Application US/10223711
; Publication No. US20030113344A1
; GENERAL INFORMATION:
;  APPLICANT: Bakaletz, Lauren O.
;  APPLICANT:  Kaumaya, Pravin T.P.
;  TITLE OF INVENTION: Synthetic Chimeric Fimbrin Peptides
;  FILE REFERENCE: 18525/04058
;  CURRENT APPLICATION NUMBER: US/10/223,711
;  CURRENT FILING DATE:  2002-08-19
;  PRIOR APPLICATION NUMBER: 09/148,711
;  PRIOR FILING DATE: 1998-09-04
;  PRIOR APPLICATION NUMBER: 08/460,502
;  PRIOR FILING DATE: 1995-06-02
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
;   LENGTH: 40
;   TYPE: PRT
;   ORGANISM: Artificial
;   FEATURE:
;   OTHER INFORMATION: synthetic construct
US-10-223-711-10

  Query Match             60.9%;  Score 106;  DB 14;  Length 40;
  Best Local Similarity   79.3%;  Pred. No. 1.1e-07;
  Matches   23;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          6 GLRPGSSGPSLKLLSEIKGVIVHRLEGVE 34
              | |    |||||||| ||||||||||||
Db         12 GTRDHKKGPSLKLLSLIKGVIVHRLEGVE 40


RESULT 4
US-09-847-102A-31
; Sequence 31, Application US/09847102A
; Publication No. US20030044409A1
; GENERAL INFORMATION:
;  APPLICANT: University of California
;  APPLICANT:  Carson, Dennis A.
;  APPLICANT:  Corr, Maripat
;  APPLICANT:  Rhee, Chae-Seo
;  APPLICANT:  Lorenzo, Leoni M.
;  APPLICANT:  Malini, Sen
;  TITLE OF INVENTION: IMMUNOLOGIC COMPOSITIONS AND METHODS FOR
;  TITLE OF INVENTION:  STUDYING AND TREATING CANCERS EXPRESSING FRIZZLED ANTIGENS
;  FILE REFERENCE: 22000-20629.00
;  CURRENT APPLICATION NUMBER: US/09/847,102A
;  CURRENT FILING DATE:  2001-05-01
;  NUMBER OF SEQ ID NOS: 138
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 31
;   LENGTH: 75
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: PFZD2-MMVF
US-09-847-102A-31

  Query Match             58.6%;  Score 102;  DB 10;  Length 75;
  Best Local Similarity   47.2%;  Pred. No. 8.3e-07;
  Matches   25;  Conservative    2;  Mismatches    2;  Indels   24;  Gaps    1;

Qy          6 GLRPGSS-----------------------GPSLKLLSEIKGVIVHRLEGVE 34
              ||:||:                       ||||||| ||||||||||||
Db         23 GLQPGAGGTPGGPGGGGAPPRYATLEHPFHCGPSLKLLSLIKGVIVHRLEGVE 75


RESULT 5
US-10-285-976-229
; Sequence 229, Application US/10285976
; Publication No. US20030165500A1
; GENERAL INFORMATION:
;  APPLICANT: Rhee, Chae-Seo
;  APPLICANT:  Malini, Sen
;  APPLICANT:  Wu, Christina
;  APPLICANT:  Leoni, Lorenzo M.
;  APPLICANT:  Corr, Maripat
;  APPLICANT:  Carson, Dennis A.
;  APPLICANT:  The Regents of the University of California
;  TITLE OF INVENTION: Wnt and Frizzled Receptors as Targets for Immunotherapy
;  TITLE OF INVENTION:  in Head and Neck Squamous Cell Carcinomas
```

```
;  FILE REFERENCE: 02307O-130320US
;  CURRENT APPLICATION NUMBER: US/10/285,976
;  CURRENT FILING DATE:  2002-11-01
;  PRIOR APPLICATION NUMBER: US 60/287,995
;  PRIOR FILING DATE: 2001-05-01
;  PRIOR APPLICATION NUMBER: WO PCT/US02/13802
;  PRIOR FILING DATE: 2002-05-01
;  NUMBER OF SEQ ID NOS: 232
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 229
;   LENGTH: 75
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:PFZD2-MMVF
;   OTHER INFORMATION: measles virus fusion (MVF) epitope fused to
;   OTHER INFORMATION: frizzled domain
US-10-285-976-229

  Query Match           58.6%;  Score 102;  DB 14;  Length 75;
  Best Local Similarity  47.2%;  Pred. No. 8.3e-07;
  Matches   25;  Conservative    2;  Mismatches    2;  Indels   24;  Gaps    1;

Qy          6 GLRPGSS-----------------------GPSLKLLSEIKGVIVHRLEGVE 34
              ||:||:                       ||||||| ||||||||||||
Db         23 GLQPGAGGTPGGPGGGGAPPRYATLEHPFHCGPSLKLLSLIKGVIVHRLEGVE 75


RESULT 6
US-10-223-711-11
; Sequence 11, Application US/10223711
; Publication No. US20030113344A1
; GENERAL INFORMATION:
;  APPLICANT: Bakaletz, Lauren O.
;  APPLICANT:  Kaumaya, Pravin T.P.
;  TITLE OF INVENTION: Synthetic Chimeric Fimbrin Peptides
;  FILE REFERENCE: 18525/04058
;  CURRENT APPLICATION NUMBER: US/10/223,711
;  CURRENT FILING DATE:  2002-08-19
;  PRIOR APPLICATION NUMBER: 09/148,711
;  PRIOR FILING DATE: 1998-09-04
;  PRIOR APPLICATION NUMBER: 08/460,502
;  PRIOR FILING DATE: 1995-06-02
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
;   LENGTH: 40
;   TYPE: PRT
;   ORGANISM: Artificial
;   FEATURE:
;   OTHER INFORMATION: synthetic construct
US-10-223-711-11

  Query Match           55.2%;  Score 96;  DB 14;  Length 40;
  Best Local Similarity  87.0%;  Pred. No. 2.8e-06;
  Matches   20;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         11 SSGPSLKLLSEIKGVIVHRLEGV 33
              ::||||||| ||||||||||||
Db         17 NTGPSLKLLSLIKGVIVHRLEGV 39


RESULT 7
US-09-848-834A-15
; Sequence 15, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
;  APPLICANT: Aphton Corporation
;  TITLE OF INVENTION: Chimeric Peptide Immunogens
;  FILE REFERENCE: 1102865-0047
;  CURRENT APPLICATION NUMBER: US/09/848,834A
;  CURRENT FILING DATE:  2001-05-04
;  PRIOR APPLICATION NUMBER: 60/202,328
;  PRIOR FILING DATE: 2000-05-05
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
;   LENGTH: 31
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the
;   OTHER INFORMATION: RH hormone linked by a spacer to amino sequence 830-844 of the
;   OTHER INFORMATION: tanus toxoid precursor (Tentoxylysin)
;   NAME/KEY: MOD_RES
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)..(10)
;   OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
;   NAME/KEY: PEPTIDE
;   LOCATION: (11)..(16)
;   OTHER INFORMATION: Spacer peptide
;   NAME/KEY: PEPTIDE
;   LOCATION: (17)..(31)
;   OTHER INFORMATION: Amino acid sequence 830-844 of the Tetanus toxoid precursor
;   OTHER INFORMATION: (Tentoxylysin)
US-09-848-834A-15

  Query Match           51.1%;  Score 89;  DB 9;  Length 31;
  Best Local Similarity  83.3%;  Pred. No. 2e-05;
  Matches   15;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 HWSYGLRPGSSGPSLKLL 19
              ||||||||||||||||: :
Db          2 HWSYGLRPGSSGPSLQYI 19


RESULT 8
US-09-848-834A-19
; Sequence 19, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
;  APPLICANT: Aphton Corporation
;  TITLE OF INVENTION: Chimeric Peptide Immunogens
;  FILE REFERENCE: 1102865-0047
;  CURRENT APPLICATION NUMBER: US/09/848,834A
;  CURRENT FILING DATE:  2001-05-04
;  PRIOR APPLICATION NUMBER: 60/202,328
;  PRIOR FILING DATE: 2000-05-05
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 19
;   LENGTH: 46
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of huma
;   OTHER INFORMATION: GnRH linked by a spacer to amino acid sequence 830-844 of Tetan
;   OTHER INFORMATION: toxoid precursor (Tentoxylysin) linked by a spacer to amino aci
;   OTHER INFORMATION: sequence 1-10 of GnRH
;   NAME/KEY: MOD_RES
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
;   NAME/KEY: MOD_RES
;   LOCATION: (46)..(46)
;   OTHER INFORMATION: Amidated glycine or glycinamide
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)..(10)
;   OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
;   NAME/KEY: PEPTIDE
;   LOCATION: (11)..(16)
;   OTHER INFORMATION: Spacer peptide
;   NAME/KEY: PEPTIDE
;   LOCATION: (17)..(31)
;   OTHER INFORMATION: Amino acid sequence 830-844 of the Tetanus toxoid precursor
```

```
;   OTHER INFORMATION: (Tentoxylysin)
;   NAME/KEY: PEPTIDE
;   LOCATION: (32)..(37)
;   OTHER INFORMATION: Spacer peptide
;   NAME/KEY: PEPTIDE
;   LOCATION: (38)..(46)
;   OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone
US-09-848-834A-19

  Query Match             51.1%;  Score 89;  DB 9;  Length 46;
  Best Local Similarity   83.3%;  Pred. No. 3.2e-05;
  Matches   15;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 HWSYGLRPGSSGPSLKLL 19
              |||||||||||||||: :
Db          2 HWSYGLRPGSSGPSLQYI 19


RESULT 9
US-09-848-834A-16
; Sequence 16, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
;  APPLICANT: Aphton Corporation
;  TITLE OF INVENTION: Chimeric Peptide Immunogens
;  FILE REFERENCE: 1102865-0047
;  CURRENT APPLICATION NUMBER: US/09/848,834A
;  CURRENT FILING DATE:  2001-05-04
;  PRIOR APPLICATION NUMBER: 60/202,328
;  PRIOR FILING DATE: 2000-05-05
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 16
;   LENGTH: 36
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the Gn
;   OTHER INFORMATION: RH hormone linked by a spacer to amino acid sequence 378-398 of
;   OTHER INFORMATION: the Plasmodium falciparum circumsporozoite (CSP) protein
;   NAME/KEY: MOD_RES
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)..(10)
;   OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
;   NAME/KEY: PEPTIDE
;   LOCATION: (11)..(16)
;   OTHER INFORMATION: Spacer peptide
;   NAME/KEY: PEPTIDE
;   LOCATION: (17)..(36)
;   OTHER INFORMATION: Amino acid sequence 378-398 of the Malaria
;   OTHER INFORMATION: (Plasmodium falciparum) circumsporozoite
;   OTHER INFORMATION: (CSP) protein
US-09-848-834A-16

  Query Match             50.0%;  Score 87;  DB 9;  Length 36;
  Best Local Similarity   100.0%;  Pred. No. 4.6e-05;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 HWSYGLRPGSSGPSL 16
              |||||||||||||||
Db          2 HWSYGLRPGSSGPSL 16


RESULT 10
US-09-848-834A-14
; Sequence 14, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
;  APPLICANT: Aphton Corporation
;  TITLE OF INVENTION: Chimeric Peptide Immunogens
;  FILE REFERENCE: 1102865-0047
;  CURRENT APPLICATION NUMBER: US/09/848,834A
;  CURRENT FILING DATE:  2001-05-04
;  PRIOR APPLICATION NUMBER: 60/202,328
;  PRIOR FILING DATE: 2000-05-05
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
;   LENGTH: 37
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the
;   OTHER INFORMATION: nRH hormone linked by a spacer to amino acid sequence 947-967 of
;   OTHER INFORMATION: the Tetanus toxoid precursor (Tentoxylysin)
;   NAME/KEY: MOD_RES
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)..(10)
;   OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
;   NAME/KEY: PEPTIDE
;   LOCATION: (11)..(16)
;   OTHER INFORMATION: Spacer peptide
;   NAME/KEY: PEPTIDE
;   LOCATION: (17)..(37)
;   OTHER INFORMATION: Amino acid sequence 947-967 of the Tetanus toxoid precursor
;   OTHER INFORMATION: (Tentoxylysin)
US-09-848-834A-14

  Query Match             50.0%;  Score 87;  DB 9;  Length 37;
  Best Local Similarity   100.0%;  Pred. No. 4.7e-05;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 HWSYGLRPGSSGPSL 16
              |||||||||||||||
Db          2 HWSYGLRPGSSGPSL 16


RESULT 11
US-09-848-834A-18
; Sequence 18, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
;  APPLICANT: Aphton Corporation
;  TITLE OF INVENTION: Chimeric Peptide Immunogens
;  FILE REFERENCE: 1102865-0047
;  CURRENT APPLICATION NUMBER: US/09/848,834A
;  CURRENT FILING DATE:  2001-05-04
;  PRIOR APPLICATION NUMBER: 60/202,328
;  PRIOR FILING DATE: 2000-05-05
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18
;   LENGTH: 50
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of humar
;   OTHER INFORMATION: GnRH linked by a spacer to amino acid sequence 947-967 of the Te
;   OTHER INFORMATION: anus toxoid precursor (Tentoxylysin) protein linked by a spacer
;   OTHER INFORMATION: o amino acid sequence 2-10 of human GnRH
;   NAME/KEY: MOD_RES
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
;   NAME/KEY: MOD_RES
;   LOCATION: (50)..(50)
;   OTHER INFORMATION: Amidated glycine or glycinamide
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)..(10)
;   OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
;   NAME/KEY: PEPTIDE
;   LOCATION: (11)..(16)
;   OTHER INFORMATION: Spacer peptide
```

```
;   NAME/KEY: PEPTIDE
;   LOCATION: (17)..(37)
;   OTHER INFORMATION: Amino acid sequence 947-967 of the Tetanus toxoid precursor (Tent
;   OTHER INFORMATION: oxylysin
;   NAME/KEY: PEPTIDE
;   LOCATION: (38)..(41)
;   OTHER INFORMATION: Spacer peptide
;   NAME/KEY: PEPTIDE
;   LOCATION: (42)..(50)
;   OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone
US-09-848-834A-18

  Query Match             50.0%;  Score 87;  DB 9;  Length 50;
  Best Local Similarity   100.0%;  Pred. No. 6.7e-05;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 HWSYGLRPGSSGPSL 16
              ||||||||||||||||
Db          2 HWSYGLRPGSSGPSL 16


RESULT 12
US-09-848-834A-20
; Sequence 20, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
;  APPLICANT: Aphton Corporation
;  TITLE OF INVENTION: Chimeric Peptide Immunogens
;  FILE REFERENCE: 1102865-0047
;  CURRENT APPLICATION NUMBER: US/09/848,834A
;  CURRENT FILING DATE:  2001-05-04
;  PRIOR APPLICATION NUMBER: 60/202,328
;  PRIOR FILING DATE: 2000-05-05
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
;   LENGTH: 51
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of human
;   OTHER INFORMATION: GnRH linked by a spacer to amino acid sequence 378-398 of Plasmod
;   OTHER INFORMATION: ium falciparum circumsporozoite (CSP) protein
;   NAME/KEY: MOD_RES
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
;   NAME/KEY: MOD_RES
;   LOCATION: (51)..(51)
;   OTHER INFORMATION: Amidated glycine or glycinamide
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)..(10)
;   OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
;   NAME/KEY: PEPTIDE
;   LOCATION: (11)..(16)
;   OTHER INFORMATION: Spacer peptide
;   NAME/KEY: PEPTIDE
;   LOCATION: (17)..(36)
;   OTHER INFORMATION: Amino acid sequence 378-398 of the Plasmodium falciparum
;   OTHER INFORMATION: circumsporozoite (CSP) protein
;   NAME/KEY: PEPTIDE
;   LOCATION: (37)..(42)
;   OTHER INFORMATION: Spacer peptide
;   NAME/KEY: PEPTIDE
;   LOCATION: (43)..(51)
;   OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone
US-09-848-834A-20

  Query Match             50.0%;  Score 87;  DB 9;  Length 51;
  Best Local Similarity   100.0%;  Pred. No. 6.9e-05;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 HWSYGLRPGSSGPSL 16
              ||||||||||||||||
Db          2 HWSYGLRPGSSGPSL 16


RESULT 13
US-09-848-834A-9
; Sequence 9, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
;  APPLICANT: Aphton Corporation
;  TITLE OF INVENTION: Chimeric Peptide Immunogens
;  FILE REFERENCE: 1102865-0047
;  CURRENT APPLICATION NUMBER: US/09/848,834A
;  CURRENT FILING DATE:  2001-05-04
;  PRIOR APPLICATION NUMBER: 60/202,328
;  PRIOR FILING DATE: 2000-05-05
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
;   LENGTH: 31
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Chimeric peptide made up of amino acid sequence 288-302 of the
;   OTHER INFORMATION: asels virus fusion protein, F linked by a spacer peptide to ami
;   OTHER INFORMATION: acid sequence 2-10 of the GnRH hormone
;   NAME/KEY: MOD_RES
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Amidated Lysine
;   NAME/KEY: PEPTIDE
;   LOCATION: (1)..(15)
;   OTHER INFORMATION: Peptide corresponds to the amino acid sequences 288-302 of the
;   OTHER INFORMATION: measles virus fusion protein, F
;   NAME/KEY: PEPTIDE
;   LOCATION: (19)..(22)
;   OTHER INFORMATION: Spacer peptide
;   NAME/KEY: PEPTIDE
;   LOCATION: (23)..(31)
;   OTHER INFORMATION: Peptide corresponds to amino acid sequences 2-10 of the human
;   OTHER INFORMATION: GnRH hormone
;   NAME/KEY: MOD_RES
;   LOCATION: (31)..(31)
;   OTHER INFORMATION: Amidated glycine or glycinamide
US-09-848-834A-9

  Query Match             49.4%;  Score 86;  DB 9;  Length 31;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-05;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         17 KLLSEIKGVIVHRLEGVE 34
              ||||||||||||||||||
Db          1 KLLSEIKGVIVHRLEGVE 18


RESULT 14
US-09-847-102A-33
; Sequence 33, Application US/09847102A
; Publication No. US20030044409A1
; GENERAL INFORMATION:
;  APPLICANT: University of California
;  APPLICANT:  Carson, Dennis A.
;  APPLICANT:  Corr, Maripat
;  APPLICANT:  Rhee, Chae-Seo
;  APPLICANT:  Lorenzo, Leoni M.
;  APPLICANT:  Malini, Sen
;  TITLE OF INVENTION: IMMUNOLOGIC COMPOSITIONS AND METHODS FOR
;  TITLE OF INVENTION:  STUDYING AND TREATING CANCERS EXPRESSING FRIZZLED ANTIGENS
;  FILE REFERENCE: 22000-20629.00
;  CURRENT APPLICATION NUMBER: US/09/847,102A
;  CURRENT FILING DATE:  2001-05-01
;  NUMBER OF SEQ ID NOS: 138
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 33
;   LENGTH: 75
```

```
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: PMMVF-FZD2
US-09-847-102A-33

  Query Match             46.0%;  Score 80;  DB 10;  Length 75;
  Best Local Similarity   89.5%;  Pred. No. 0.001;
  Matches   17;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         16 LKLLSEIKGVIVHRLEGVE 34
              :|||| ||||||||||||||
Db          1 MKLLSLIKGVIVHRLEGVE 19


RESULT 15
US-10-285-976-231
; Sequence 231, Application US/10285976
; Publication No. US20030165500A1
; GENERAL INFORMATION:
;  APPLICANT: Rhee, Chae-Seo
;  APPLICANT:  Malini, Sen
;  APPLICANT:  Wu, Christina
;  APPLICANT:  Leoni, Lorenzo M.
;  APPLICANT:  Corr, Maripat
;  APPLICANT:  Carson, Dennis A.
;  APPLICANT:  The Regents of the University of California
;  TITLE OF INVENTION: Wnt and Frizzled Receptors as Targets for Immunotherapy
;  TITLE OF INVENTION:  in Head and Neck Squamous Cell Carcinomas
;  FILE REFERENCE: 02307O-130320US
;  CURRENT APPLICATION NUMBER: US/10/285,976
;  CURRENT FILING DATE:  2002-11-01
;  PRIOR APPLICATION NUMBER: US 60/287,995
;  PRIOR FILING DATE: 2001-05-01
;  PRIOR APPLICATION NUMBER: WO PCT/US02/13802
;  PRIOR FILING DATE: 2002-05-01
;  NUMBER OF SEQ ID NOS: 232
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 231
;   LENGTH: 75
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:PMMVF-ZD2
;   OTHER INFORMATION: measles virus fusion (MVF) epitope fused to
;   OTHER INFORMATION: frizzled domain
US-10-285-976-231

  Query Match             46.0%;  Score 80;  DB 14;  Length 75;
  Best Local Similarity   89.5%;  Pred. No. 0.001;
  Matches   17;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         16 LKLLSEIKGVIVHRLEGVE 34
              :|||| ||||||||||||||
Db          1 MKLLSLIKGVIVHRLEGVE 19


Search completed: March 10, 2004, 10:25:48
Job time : 26.7237 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:48 ; Search time 51.1984 Seconds
(without alignments)
187.635 Million cell updates/sec

Title: US-09-848-834A-13
Perfect score: 174
Sequence: 1 XHWSYGLRPGSSGPSLKLLSEIKGVIVHRLEGVE 34

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	173	99.4	34	5	AAU11424	Aau11424 Synthetic
2	173	99.4	47	5	AAU11428	Aau11428 Synthetic
3	106	60.9	40	2	AAW67581	Aaw67581 Synthetic
4	106	60.9	40	3	AAY79986	Aay79986 Measles v
5	106	60.9	40	6	ADA25172	Ada25172 Chimeric
6	106	60.9	40	7	ADC89661	Adc89661 H. influe
7	102	58.6	75	6	ABP72235	Abp72235 Frizzled
8	101	58.0	40	2	AAW67582	Aaw67582 Synthetic
9	96	55.2	40	6	ADA25173	Ada25173 Chimeric
10	96	55.2	40	7	ADC89662	Adc89662 H. influe
11	89	51.1	31	5	AAU11426	Aau11426 Synthetic
12	89	51.1	46	5	AAU11430	Aau11430 Synthetic
13	87	50.0	36	5	AAU11427	Aau11427 Synthetic
14	87	50.0	37	5	AAU11425	Aau11425 Synthetic
15	87	50.0	50	5	AAU11429	Aau11429 Synthetic
16	87	50.0	51	5	AAU11431	Aau11431 Synthetic
17	86	49.4	31	5	AAU11420	Aau11420 Synthetic
18	80	46.0	75	6	ABP72236	Abp72236 Frizzled
19	78	44.8	18	2	AAW35441	Aaw35441 T-cell st
20	77	44.3	16	5	AAU11419	Aau11419 Mesles vi
21	77	44.3	25	2	AAR62705	Aar62705 LHRH-cont
22	74	42.5	20	2	AAW57161	Aaw57161 Measles v
23	74	42.5	72	2	AAR88391	Aar88391 Measles v
24	73.5	42.2	34	2	AAW05620	Aaw05620 mIgE2-GG-
25	73.5	42.2	43	2	AAW05618	Aaw05618 mIgE1-GG-
26	73	42.0	42	2	AAR62708	Aar62708 LHRH-cont
27	72	41.4	15	2	AAR62697	Aar62697 Helper T
28	72	41.4	15	2	AAR82082	Aar82082 Measles v
29	72	41.4	15	2	AAR82591	Aar82591 Measles v
30	72	41.4	15	2	AAR88401	Aar88401 Measles v
31	72	41.4	15	2	AAR88392	Aar88392 Measles v
32	72	41.4	15	2	AAW05604	Aaw05604 Measles v
33	72	41.4	15	3	AAY68540	Aay68540 Helper T
34	72	41.4	15	3	AAY91121	Aay91121 Measles v
35	72	41.4	15	3	AAY58764	Aay58764 Measles v
36	72	41.4	15	3	AAY80054	Aay80054 Pathogen
37	72	41.4	15	3	AAY54537	Aay54537 T helper
38	72	41.4	15	3	AAY44762	Aay44762 Measles v
39	72	41.4	15	4	AAB84440	Aab84440 Amino aci
40	72	41.4	15	4	AAB68638	Aab68638 HER-2 B c
41	72	41.4	15	5	AAU97873	Aau97873 Measles v
42	72	41.4	15	5	ABG68169	Abg68169 Pathogen-
43	72	41.4	15	5	ABG68183	Abg68183 Measles v
44	72	41.4	15	6	AAE35614	Aae35614 Measles v
45	72	41.4	15	6	AAE35628	Aae35628 Measles v

ALIGNMENTS

RESULT 1
AAU11424
ID AAU11424 standard; peptide; 34 AA.
XX
AC AAU11424;
XX
DT 12-MAR-2002 (first entry)
XX
DE Synthetic immunogen peptide 5.
XX
KW Gonadotrophin releasing hormone; GnRH; synthetic immunogen;
KW luteinising hormone releasing hormone; LHRH; contraceptive;
KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;
KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.
XX
OS Measles virus.
OS Mammalia.
OS Synthetic.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT Peptide 1. .10
FT /note= "Gonadotrophin releasing hormone epitope"
FT Misc-difference 1
FT /label= OTHER
FT /note= "Other= Pyro-glutamic acid or 5-oxo proline"
FT Peptide 11. .16
FT /note= "Spacer peptide"
FT Peptide 17. .34
FT /note= "Measles virus fusion protein F epitope"
XX
PN WO200185763-A2.
XX
PD 15-NOV-2001.
XX
PF 04-MAY-2001; 2001WO-US014363.
XX
PR 05-MAY-2000; 2000US-0202328P.
XX
PA (APHT-) APHTON CORP.
XX
PI Grimes S, Michaeli D, Stevens VC;
XX
DR WPI; 2002-049440/06.
XX
PT Novel synthetic immunogen for inducing immune response against
PT gonadotropin releasing hormone, comprises fusion peptide having

PT promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
PT or its analog.
XX
PS Claim 11; Page 9; 43pp; English.
XX
CC The invention relates to a synthetic immunogen for inducing specific
CC antibodies against gonadotropin releasing hormone (GnRH also known as
CC luteinising hormone releasing hormone, LHRH) comprising a fusion peptide
CC which comprises a promiscuous helper T-cell peptide epitope and
CC immunomimic peptide epitope or its analogue. The synthetic immunogen is
CC useful inducing an immune response against GnRH in an animal subject, and
CC as such is useful as a contraceptive and in the treatment of diseases
CC such as cancer (of the breast, uterus and other gynaecological cancer),
CC endometriosis, uterine fibroids, benign prostatic hypertrophy and
CC prostate cancer. The immunogen is effective in eliciting high and
CC specific anti-GnRH antibody titres. The present sequence is a synthetic
CC immunogen of the invention
XX
SQ Sequence 34 AA;

Query Match 99.4%; Score 173; DB 5; Length 34;
Best Local Similarity 100.0%; Pred. No. 7.8e-18;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HWSYGLRPGSSGPSLKLLSEIKGVIVHRLEGVE 34
||||||||||||||||||||||||||||||||
Db 2 HWSYGLRPGSSGPSLKLLSEIKGVIVHRLEGVE 34

RESULT 2
AAU11428
ID AAU11428 standard; peptide; 47 AA.
XX
AC AAU11428;
XX
DT 12-MAR-2002 (first entry)
XX
DE Synthetic immunogen peptide 9.
XX
KW Gonadotrophin releasing hormone; GnRH; synthetic immunogen;
KW luteinising hormone releasing hormone; LHRH; contraceptive;
KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;
KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.
XX
OS Plasmodium falciparum.
OS Mammalia.
OS Synthetic.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT Peptide 1. .10
FT /note= "Gonadotrophin releasing hormone epitope (1. .10
FT aa)"
FT Misc-difference 1
FT /label= OTHER
FT /note= "Other= Pyro-glutamic acid or 5-oxo proline"
FT Peptide 11. .16
FT /note= "Spacer peptide"
FT Peptide 17. .34
FT /note= "Malaria CSP protein (288-302 aa)"
FT Peptide 35. .38
FT /note= "Spacer peptide"
FT Peptide 39. .47
FT /note= "Gonadotrophin releasing hormone epitope (2-10
FT aa)"
FT Modified-site 47
FT /note= "Amidated glycine or glycinamide"
XX
PN WO200185763-A2.
XX
PD 15-NOV-2001.
XX
PF 04-MAY-2001; 2001WO-US014363.
XX
PR 05-MAY-2000; 2000US-0202328P.
XX
PA (APHT-) APHTON CORP.
XX
PI Grimes S, Michaeli D, Stevens VC;
XX
DR WPI; 2002-049440/06.
XX
PT Novel synthetic immunogen for inducing immune response against
PT gonadotropin releasing hormone, comprises fusion peptide having
PT promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
PT or its analog.
XX
PS Claim 11; Page 11; 43pp; English.
XX
CC The invention relates to a synthetic immunogen for inducing specific
CC antibodies against gonadotropin releasing hormone (GnRH also known as
CC luteinising hormone releasing hormone, LHRH) comprising a fusion peptide
CC which comprises a promiscuous helper T-cell peptide epitope and
CC immunomimic peptide epitope or its analogue. The synthetic immunogen is
CC useful inducing an immune response against GnRH in an animal subject, and
CC as such is useful as a contraceptive and in the treatment of diseases
CC such as cancer (of the breast, uterus and other gynaecological cancer),
CC endometriosis, uterine fibroids, benign prostatic hypertrophy and
CC prostate cancer. The immunogen is effective in eliciting high and
CC specific anti-GnRH antibody titres. The present sequence is a synthetic
CC immunogen of the invention
XX
SQ Sequence 47 AA;

Query Match 99.4%; Score 173; DB 5; Length 47;
Best Local Similarity 100.0%; Pred. No. 1.2e-17;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HWSYGLRPGSSGPSLKLLSEIKGVIVHRLEGVE 34
||||||||||||||||||||||||||||||||
Db 2 HWSYGLRPGSSGPSLKLLSEIKGVIVHRLEGVE 34

RESULT 3
AAW67581
ID AAW67581 standard; peptide; 40 AA.
XX
AC AAW67581;
XX
DT 02-MAR-1999 (first entry)
XX
DE Synthetic chimer fimbrin/T-cell epitope peptide LB1.
XX
KW Chimeric; non-typable Haemophilus influenzae; fimbrin; T-cell epitope;
KW immunogenic composition; immune response.
XX
OS Synthetic.
XX
PN US5843464-A.
XX
PD 01-DEC-1998.
XX
PF 02-JUN-1995; 95US-00460502.
XX
PR 02-JUN-1995; 95US-00460502.
XX
PA (OHIS) UNIV OHIO STATE.
XX
PI Kaumaya PTP, Bakaletz LO;
XX
DR WPI; 1999-044514/04.
XX
PT Synthetic chimeric fimbrin peptide - useful for vaccination against non-
PT typable Haemophilus influenzae.
XX

PS Claim 4; Col 4; 16pp; English.
XX
CC The invention relates to the manufacture of a synthetic chimeric peptide
CC comprising a non-typable Haemophilus influenzae fimbrin peptide fused via
CC a linker peptide to a T-cell epitope peptide. The chimeric peptide is
CC used in immunogenic compositions which induce an immune response against
CC non-typable Haemophilus influenzae. This sequence represents an example
CC of a chimeric fimbrin/T-cell epitope peptide and is designated LB1. The
CC peptide comprises a 19 amino acid sequence corresponding to amino acids
CC 117-135 of the fimbrin protein, the linker sequence and amino acid 288-
CC 302 of the measles virus fusion protein (a T-cell epitope)
XX
SQ Sequence 40 AA;

Query Match 60.9%; Score 106; DB 2; Length 40;
Best Local Similarity 79.3%; Pred. No. 6e-08;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 6 GLRPGSSGPSLKLLSEIKGVIVHRLEGVE 34
Db 12 GTRDHKKGPSLKLLSLIKGVIVHRLEGVE 40

RESULT 4
AAY79986
ID AAY79986 standard; peptide; 40 AA.
XX
AC AAY79986;
XX
DT 15-MAY-2000 (first entry)
XX
DE Measles virus fusion protein T-cell promiscuous epitope.
XX
KW Vaccine; non-typeable Haemophilus influenzae; ntHi; infection;
KW chimeric protein; Haemophilus influenzae; P5-like fimbrin protein;
KW lipoprotein D; LB1(f); immunogenic; antigenic; otitis media; sinusitis;
KW conjunctivitis; lower respiratory tract infection.
XX
OS Measles virus.
OS Synthetic.
XX
PN WO9964067-A2.
XX
PD 16-DEC-1999.
XX
PF 28-MAY-1999; 99WO-US011980.
XX
PR 11-JUN-1998; 98GB-00012613.
XX
PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
PA (OHIS) UNIV OHIO STATE RES FOUND.
XX
PI Bakaletz LO, Cohen J, Dequesne G, Lobet Y;
XX
DR WPI; 2000-116457/10.
XX
PT Novel antigenic P5-like fimbrin subunit peptides used in vaccines against
PT Haemophilus influenza.
XX
PS Example 4; Page 38; 68pp; English.
XX
CC The present invention describes antigenic P5-like fimbrin subunit
CC peptides (LB1(f) peptides) of P5-like fimbrin proteins from various
CC Haemophilus influenzae strains. The peptides are used for diagnosis,
CC prevention, and treatment of Haemophilus influenzae infections, such as
CC otitis media, sinusitis, conjunctivitis, or lower respiratory tract
CC infection. The peptides may also be used in vaccines against H.
CC influenzae. Antibodies and probes from the present invention can be used
CC for diagnosis of H. influenzae infection. AAY79955 to AAY79993, and
CC AAZ91201 to AAZ91252, represent sequences used in the exemplification of
CC the present invention
XX
SQ Sequence 40 AA;

Query Match 60.9%; Score 106; DB 3; Length 40;
Best Local Similarity 79.3%; Pred. No. 6e-08;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 6 GLRPGSSGPSLKLLSEIKGVIVHRLEGVE 34
Db 12 GTRDHKKGPSLKLLSLIKGVIVHRLEGVE 40

RESULT 5
ADA25172
ID ADA25172 standard; peptide; 40 AA.
XX
AC ADA25172;
XX
DT 20-NOV-2003 (first entry)
XX
DE Chimeric fimbrin peptide LB1.
XX
KW fimbrin; non-typable Haemophilus influenzae; NTHi infection;
KW otitis media.
XX
OS Chimeric.
OS Synthetic.
OS Haemophilus influenzae.
OS Measles virus.
XX
PN US6436405-B1.
XX
PD 20-AUG-2002.
XX
PF 04-SEP-1998; 98US-00148711.
XX
PR 02-JUN-1995; 95US-00460502.
XX
PA (OHIS) UNIV OHIO STATE.
XX
PI Bakaletz LO, Kaumaya PTP;
XX
DR WPI; 2003-615247/58.
XX
PT Synthetic chimeric fimbrin peptide, useful for treating Haemophilus
PT influenzae infections.
XX
PS Claim 10; Col 4; 16pp; English.
XX
CC The invention relates to a synthetic chimeric fimbrin peptide. The
CC peptide is useful for treating a non-typable Haemophilus influenzae
CC (NTHi) infection and otitis media. The synthetic peptides do not require
CC tedious purification techniques. The present sequence represents the
CC amino acid sequence of the chimeric fimbrin peptide LB1.
XX
SQ Sequence 40 AA;

Query Match 60.9%; Score 106; DB 6; Length 40;
Best Local Similarity 79.3%; Pred. No. 6e-08;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 6 GLRPGSSGPSLKLLSEIKGVIVHRLEGVE 34
Db 12 GTRDHKKGPSLKLLSLIKGVIVHRLEGVE 40

RESULT 6
ADC89661
ID ADC89661 standard; peptide; 40 AA.
XX
AC ADC89661;
XX
DT 01-JAN-2004 (first entry)
XX
DE H. influenzae fimbrin peptide/T cell epitope chimaera LB1.

XX
KW Fimbrin; T cell epitope; vaccine; otitis media; auditory;
KW antiinflammatory; LB1.
XX
OS Chimeric.
OS Haemophilus influenzae.
OS Measles virus.
XX
PN US2003113344-A1.
XX
PD 19-JUN-2003.
XX
PF 19-AUG-2002; 2002US-00223711.
XX
PR 04-SEP-1998; 98US-00148711.
XX
PA (BAKA/) BAKALETZ L O.
PA (KAUM/) KAUMAYA P T P.
XX
PI Bakaletz LO, Kaumaya PTP;
XX
DR WPI; 2003-810881/76.
XX
PT Novel synthetic chimeric fimbrin peptide LB1 or LB2 comprising a first
PT peptide unit, T cell epitope as second peptide unit and third linker
PT peptide unit, useful for preventing or reducing severity of otitis media.
XX
PS Claim 8; SEQ ID NO 10; 15pp; English.
XX
CC The invention relates to a synthetic chimaeric fimbrin peptide LB1 or LB2
CC comprises a first peptide unit derived from H. influenzae fimbrin , a
CC second peptide unit containing a T cell epitope and a third linker
CC peptide which connects the first peptide to the second. The chimaeric
CC peptide is useful for inducing an immune response in animals against non-
CC typable Haemophilus influenzae (NTHi) and for preventing or reducing
CC adherence of NTHi to host cells thereby preventing or reducing the
CC severity of otitis media. The present sequence is an H. influenzae
CC fimbrin peptide/measles virus T cell epitope chimaeric peptide of the
CC invention, LB1.
XX
SQ Sequence 40 AA;

Query Match 60.9%; Score 106; DB 7; Length 40;
Best Local Similarity 79.3%; Pred. No. 6e-08;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy          6 GLRPGSSGPSLKLLSEIKGVIVHRLEGVE 34
              ||    |||||||| ||||||||||||
Db         12 GTRDHKKGPSLKLLSLIKGVIVHRLEGVE 40
```

RESULT 7
ABP72235
ID ABP72235 standard; protein; 75 AA.
XX
AC ABP72235;
XX
DT 28-APR-2003 (first entry)
XX
DE Frizzled putative B-cell epitope-measles MVF epitope fusion.
XX
KW Frizzled; Fzd-2; receptor; human; head and neck squamous carcinoma;
KW cancer; diagnosis; immunotherapy; gene therapy; cytostatic; antitumour;
KW epitope.
XX
OS Homo sapiens.
OS Measles virus.
OS Synthetic.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT Peptide 1. .53
FT /note= "Frizzled putative B cell epitope"
FT Peptide 54. .57
FT /note= "Linker peptide"
FT Protein 58. .75
FT /note= "Measles MVF T cell epitope"
XX
PN WO200288081-A2.
XX
PD 07-NOV-2002.
XX
PF 01-MAY-2002; 2002WO-US013802.
XX
PR 01-MAY-2001; 2001US-0287995P.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Rhee C, Sen M, Wu C, Leoni LM, Corr M, Carson DA;
XX
DR WPI; 2003-111859/10.
DR N-PSDB; ABZ58287.
XX
PT Determining overexpression of wnt or frizzled gene in a tumor cell useful
PT for immunotherapy in head and neck squamous cell carcinoma comprises
PT reverse transcription PCR with primers directed against the wnt or
PT frizzled gene.
XX
PS Example 1; Page 22; 51pp; English.
XX
CC The present sequence is that of a fusion protein in which a putative B
CC cell epitope of human frizzled Fzd-2 is joined via a peptide linker to
CC the measles virus MVF T helper epitope. In an example from the invention,
CC DNA encoding the construct was inserted into a plasmid vector, and mice
CC were injected with plasmid DNA and then boosted with peptide or
CC recombinant protein. The B-cell epitope may need to be redesigned to
CC eliminate cross-reactivity in the humoral response to other frizzled
CC isoforms. The invention is based on the finding that most head and neck
CC squamous carcinoma cell lines (HNSCC) overexpress one or more wingless
CC (Wnt) or frizzled (Fzd) mRNAs, making the Wnt and Fzd receptors useful
CC targets for immunotherapy of this common cancer. The invention provides
CC claimed methods for determining, by RT-PCR, overexpression of a Wnt or
CC Fzd gene in a tumour cell, for detecting overexpression of a Wnt and/or
CC Fzd protein in a cancer, and for altering the growth of a cell
CC overexpressing a Wnt and/or Fzd protein by contacting the cell with an
CC antibody against the protein, or with a synthetic peptide, recombinant
CC protein or DNA vector comprising a non-homologous region of Wnt and/or
CC Fzd proteins
XX
SQ Sequence 75 AA;

Query Match 58.6%; Score 102; DB 6; Length 75;
Best Local Similarity 47.2%; Pred. No. 5e-07;
Matches 25; Conservative 2; Mismatches 2; Indels 24; Gaps 1;

```
Qy          6 GLRPGSS------------------------GPSLKLLSEIKGVIVHRLEGVE 34
                ||:||:                        |||||||| ||||||||||||
Db         23 GLQPGAGGTPGGPGGGGAPPRYATLEHPFHCGPSLKLLSLIKGVIVHRLEGVE 75
```

RESULT 8
AAW67582
ID AAW67582 standard; peptide; 40 AA.
XX
AC AAW67582;
XX
DT 02-MAR-1999 (first entry)
XX
DE Synthetic chimer fimbrin/T-cell epitope peptide LB2.
XX
KW Chimeric; non-typable Haemophilus influenzae; fimbrin; T-cell epitope;
KW immunogenic composition; immune response.
XX
OS Synthetic.
XX
PN US5843464-A.

```
XX
PD   01-DEC-1998.
XX
PF   02-JUN-1995;   95US-00460502.
XX
PR   02-JUN-1995;   95US-00460502.
XX
PA   (OHIS ) UNIV OHIO STATE.
XX
PI   Kaumaya PTP,  Bakaletz LO;
XX
DR   WPI; 1999-044514/04.
XX
PT   Synthetic chimeric fimbrin peptide - useful for vaccination against non-
PT   typable Haemophilus influenzae.
XX
PS   Disclosure; Col 4; 16pp; English.
XX
CC   The invention relates to the manufacture of a synthetic chimeric peptide
CC   comprising a non-typable Haemophilus influenzae fimbrin peptide fused via
CC   a linker peptide to a T-cell epitope peptide. The chimeric peptide is
CC   used in immunogenic compositions which induce an immune response against
CC   non-typable Haemophilus influenzae. This sequence represents an example
CC   of a chimeric fimbrin/T-cell epitope peptide and is designated LB2. The
CC   peptide comprises a 18 amino acid sequence corresponding to amino acids
CC   163-180 of the fimbrin protein, the linker sequence and amino acid 288-
CC   302 of the measles virus fusion protein (a T-cell epitope)
XX
SQ   Sequence 40 AA;

  Query Match             58.0%;  Score 101;  DB 2;  Length 40;
  Best Local Similarity   87.5%;  Pred. No. 3.3e-07;
  Matches   21;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         11 SSGPSLKLLSEIKGVIVHRLEGVE 34
              ::|||||||| ||||||||||||
Db         17 NTGPSLKLLSLIKGVIVHRLEGVE 40


RESULT 9
ADA25173
ID   ADA25173 standard; peptide; 40 AA.
XX
AC   ADA25173;
XX
DT   20-NOV-2003  (first entry)
XX
DE   Chimeric fimbrin peptide LB2.
XX
KW   fimbrin; non-typable Haemophilus influenzae; NTHi infection;
KW   otitis media.
XX
OS   Chimeric.
OS   Synthetic.
OS   Haemophilus influenzae.
OS   Measles virus.
XX
PN   US6436405-B1.
XX
PD   20-AUG-2002.
XX
PF   04-SEP-1998;   98US-00148711.
XX
PR   02-JUN-1995;   95US-00460502.
XX
PA   (OHIS ) UNIV OHIO STATE.
XX
PI   Bakaletz LO,  Kaumaya PTP;
XX
DR   WPI; 2003-615247/58.
XX
PT   Synthetic chimeric fimbrin peptide, useful for treating Haemophilus
PT   influenzae infections.
XX
PS   Claim 5; Col 4; 16pp; English.
XX
CC   The invention relates to a synthetic chimeric fimbrin peptide. The
CC   peptide is useful for treating a non-typable Haemophilus influenzae
CC   (NTHi) infection and otitis media. The synthetic peptides do not require
CC   tedious purification techniques. The present sequence represents the
CC   amino acid sequence of the chimeric fimbrin peptide LB2.
XX
SQ   Sequence 40 AA;

  Query Match             55.2%;  Score 96;  DB 6;  Length 40;
  Best Local Similarity   87.0%;  Pred. No. 1.8e-06;
  Matches   20;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         11 SSGPSLKLLSEIKGVIVHRLEGV 33
              ::|||||||| |||||||||||
Db         17 NTGPSLKLLSLIKGVIVHRLEGV 39


RESULT 10
ADC89662
ID   ADC89662 standard; peptide; 40 AA.
XX
AC   ADC89662;
XX
DT   01-JAN-2004  (first entry)
XX
DE   H. influenzae fimbrin peptide/T cell epitope chimaera LB2.
XX
KW   Fimbrin; T cell epitope; vaccine; otitis media; auditory;
KW   antiinflammatory; LB2.
XX
OS   Chimeric.
OS   Haemophilus influenzae.
OS   Measles virus.
XX
PN   US2003113344-A1.
XX
PD   19-JUN-2003.
XX
PF   19-AUG-2002; 2002US-00223711.
XX
PR   04-SEP-1998;   98US-00148711.
XX
PA   (BAKA/) BAKALETZ L O.
PA   (KAUM/) KAUMAYA P T P.
XX
PI   Bakaletz LO,  Kaumaya PTP;
XX
DR   WPI; 2003-810881/76.
XX
PT   Novel synthetic chimeric fimbrin peptide LB1 or LB2 comprising a first
PT   peptide unit, T cell epitope as second peptide unit and third linker
PT   peptide unit, useful for preventing or reducing severity of otitis media.
XX
PS   Claim 9; SEQ ID NO 11; 15pp; English.
XX
CC   The invention relates to a synthetic chimaeric fimbrin peptide LB1 or LB2
CC   comprises a first peptide unit derived from H. influenzae fimbrin , a
CC   second peptide unit containing a T cell epitope and a third linker
CC   peptide which connects the first peptide to the second. The chimaeric
CC   peptide is useful for inducing an immune response in animals against non-
CC   typable Haemophilus influenzae (NTHi) and for preventing or reducing
CC   adherence of NTHi to host cells thereby preventing or reducing the
CC   severity of otitis media. The present sequence is an H. influenzae
CC   fimbrin peptide/measles virus T cell epitope chimaeric peptide of the
CC   invention, LB2.
XX
SQ   Sequence 40 AA;

  Query Match             55.2%;  Score 96;  DB 7;  Length 40;
  Best Local Similarity   87.0%;  Pred. No. 1.8e-06;
```

```
  Matches   20;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         11 SSGPSLKLLSEIKGVIVHRLEGV 33
              ::|||||||| ||||||||||||
Db         17 NTGPSLKLLSLIKGVIVHRLEGV 39


RESULT 11
AAU11426
ID   AAU11426 standard; peptide; 31 AA.
XX
AC   AAU11426;
XX
DT   12-MAR-2002  (first entry)
XX
DE   Synthetic immunogen peptide 7.
XX
KW   Gonadotrophin releasing hormone; GnRH; synthetic immunogen;
KW   luteinising hormone releasing hormone; LHRH; contraceptive;
KW   promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
KW   breast cancer; uterine cancer; gynaecological cancer; endometriosis;
KW   uterine fibroid; benign prostatic hypertrophy; prostate cancer.
XX
OS   Clostridium tetani.
OS   Mammalia.
OS   Synthetic.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .10
FT                   /note= "Gonadotrophin releasing hormone epitope"
FT   Misc-difference 1
FT                   /label= OTHER
FT                   /note= "Other= Pyro-glutamic acid or 5-oxo proline"
FT   Peptide         11. .16
FT                   /note= "Spacer peptide"
FT   Peptide         17. .31
FT                   /note= "Tetanus toxoid sequence (830-844 aa)"
XX
PN   WO200185763-A2.
XX
PD   15-NOV-2001.
XX
PF   04-MAY-2001; 2001WO-US014363.
XX
PR   05-MAY-2000; 2000US-0202328P.
XX
PA   (APHT-) APHTON CORP.
XX
PI   Grimes S,  Michaeli D,  Stevens VC;
XX
DR   WPI; 2002-049440/06.
XX
PT   Novel synthetic immunogen for inducing immune response against
PT   gonadotropin releasing hormone, comprises fusion peptide having
PT   promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
PT   or its analog.
XX
PS   Claim 11; Page 10; 43pp; English.
XX
CC   The invention relates to a synthetic immunogen for inducing specific
CC   antibodies against gonadotropin releasing hormone (GnRH also known as
CC   luteinising hormone releasing hormone, LHRH) comprising a fusion peptide
CC   which comprises a promiscuous helper T-cell peptide epitope and
CC   immunomimic peptide epitope or its analogue. The synthetic immunogen is
CC   useful inducing an immune response against GnRH in an animal subject, and
CC   as such is useful as a contraceptive and in the treatment of diseases
CC   such as cancer (of the breast, uterus and other gynaecological cancer),
CC   endometriosis, uterine fibroids, benign prostatic hypertrophy and
CC   prostate cancer. The immunogen is effective in eliciting high and
CC   specific anti-GnRH antibody titres. The present sequence is a synthetic
CC   immunogen of the invention
XX
SQ   Sequence 31 AA;

  Query Match             51.1%;  Score 89;  DB 5;  Length 31;
  Best Local Similarity   83.3%;  Pred. No. 1.4e-05;
  Matches   15;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 HWSYGLRPGSSGPSLKLL 19
              ||||||||||||||: :
Db          2 HWSYGLRPGSSGPSLQYI 19


RESULT 12
AAU11430
ID   AAU11430 standard; peptide; 46 AA.
XX
AC   AAU11430;
XX
DT   12-MAR-2002  (first entry)
XX
DE   Synthetic immunogen peptide 11.
XX
KW   Gonadotrophin releasing hormone; GnRH; synthetic immunogen;
KW   luteinising hormone releasing hormone; LHRH; contraceptive;
KW   promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
KW   breast cancer; uterine cancer; gynaecological cancer; endometriosis;
KW   uterine fibroid; benign prostatic hypertrophy; prostate cancer.
XX
OS   Clostridium tetani.
OS   Mammalia.
OS   Synthetic.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .10
FT                   /note= "Gonadotrophin releasing hormone epitope (1. .10
FT                   aa)"
FT   Misc-difference 1
FT                   /label= OTHER
FT                   /note= "Other= Pyro-glutamic acid or 5-oxo proline"
FT   Peptide         11. .16
FT                   /note= "Spacer peptide"
FT   Peptide         17. .31
FT                   /note= "Tetanus toxoid (830-844 aa)"
FT   Peptide         32. .37
FT                   /note= "Spacer peptide"
FT   Peptide         38. .46
FT                   /note= "Gonadotrophin releasing hormone epitope (2-10
FT                   aa)"
FT   Modified-site   46
FT                   /note= "Amidated glycine or glycinamide"
XX
PN   WO200185763-A2.
XX
PD   15-NOV-2001.
XX
PF   04-MAY-2001; 2001WO-US014363.
XX
PR   05-MAY-2000; 2000US-0202328P.
XX
PA   (APHT-) APHTON CORP.
XX
PI   Grimes S,  Michaeli D,  Stevens VC;
XX
DR   WPI; 2002-049440/06.
XX
PT   Novel synthetic immunogen for inducing immune response against
PT   gonadotropin releasing hormone, comprises fusion peptide having
PT   promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
PT   or its analog.
XX
PS   Claim 11; Page 12; 43pp; English.
XX
CC   The invention relates to a synthetic immunogen for inducing specific
```

```
CC   antibodies against gonadotropin releasing hormone (GnRH also known as
CC   luteinising hormone releasing hormone, LHRH) comprising a fusion peptide
CC   which comprises a promiscuous helper T-cell peptide epitope and
CC   immunomimic peptide epitope or its analogue. The synthetic immunogen is
CC   useful inducing an immune response against GnRH in an animal subject, and
CC   as such is useful as a contraceptive and in the treatment of diseases
CC   such as cancer (of the breast, uterus and other gynaecological cancer),
CC   endometriosis, uterine fibroids, benign prostatic hypertrophy and
CC   prostate cancer. The immunogen is effective in eliciting high and
CC   specific anti-GnRH antibody titres. The present sequence is a synthetic
CC   immunogen of the invention
XX
SQ   Sequence 46 AA;

  Query Match             51.1%;  Score 89;  DB 5;  Length 46;
  Best Local Similarity   83.3%;  Pred. No. 2.2e-05;
  Matches   15;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 HWSYGLRPGSSGPSLKLL 19
              ||||||||||||||||:. :
Db          2 HWSYGLRPGSSGPSLQYI 19


RESULT 13
AAU11427
ID   AAU11427 standard; peptide; 36 AA.
XX
AC   AAU11427;
XX
DT   12-MAR-2002  (first entry)
XX
DE   Synthetic immunogen peptide 8.
XX
KW   Gonadotrophin releasing hormone; GnRH; synthetic immunogen;
KW   luteinising hormone releasing hormone; LHRH; contraceptive;
KW   promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
KW   breast cancer; uterine cancer; gynaecological cancer; endometriosis;
KW   uterine fibroid; benign prostatic hypertrophy; prostate cancer.
XX
OS   Plasmodium falciparum.
OS   Mammalia.
OS   Synthetic.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .10
FT                   /note= "Gonadotrophin releasing hormone epitope"
FT   Misc-difference 1
FT                   /label= OTHER
FT                   /note= "Other= Pyro-glutamic acid or 5-oxo proline"
FT   Peptide         11. .16
FT                   /note= "Spacer peptide"
FT   Peptide         17. .36
FT                   /note= "Malaria CSP protein (378-398 aa)"
XX
PN   WO200185763-A2.
XX
PD   15-NOV-2001.
XX
PF   04-MAY-2001; 2001WO-US014363.
XX
PR   05-MAY-2000; 2000US-0202328P.
XX
PA   (APHT-) APHTON CORP.
XX
PI   Grimes S,  Michaeli D,  Stevens VC;
XX
DR   WPI; 2002-049440/06.
XX
PT   Novel synthetic immunogen for inducing immune response against
PT   gonadotropin releasing hormone, comprises fusion peptide having
PT   promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
PT   or its analog.
XX
PS   Claim 11; Page 10; 43pp; English.
XX
CC   The invention relates to a synthetic immunogen for inducing specific
CC   antibodies against gonadotropin releasing hormone (GnRH also known as
CC   luteinising hormone releasing hormone, LHRH) comprising a fusion peptide
CC   which comprises a promiscuous helper T-cell peptide epitope and
CC   immunomimic peptide epitope or its analogue. The synthetic immunogen is
CC   useful inducing an immune response against GnRH in an animal subject, and
CC   as such is useful as a contraceptive and in the treatment of diseases
CC   such as cancer (of the breast, uterus and other gynaecological cancer),
CC   endometriosis, uterine fibroids, benign prostatic hypertrophy and
CC   prostate cancer. The immunogen is effective in eliciting high and
CC   specific anti-GnRH antibody titres. The present sequence is a synthetic
CC   immunogen of the invention
XX
SQ   Sequence 36 AA;

  Query Match             50.0%;  Score 87;  DB 5;  Length 36;
  Best Local Similarity   100.0%;  Pred. No. 3.2e-05;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 HWSYGLRPGSSGPSL 16
              |||||||||||||||
Db          2 HWSYGLRPGSSGPSL 16


RESULT 14
AAU11425
ID   AAU11425 standard; peptide; 37 AA.
XX
AC   AAU11425;
XX
DT   12-MAR-2002  (first entry)
XX
DE   Synthetic immunogen peptide 6.
XX
KW   Gonadotrophin releasing hormone; GnRH; synthetic immunogen;
KW   luteinising hormone releasing hormone; LHRH; contraceptive;
KW   promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
KW   breast cancer; uterine cancer; gynaecological cancer; endometriosis;
KW   uterine fibroid; benign prostatic hypertrophy; prostate cancer.
XX
OS   Clostridium tetani.
OS   Mammalia.
OS   Synthetic.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .10
FT                   /note= "Gonadotrophin releasing hormone epitope"
FT   Misc-difference 1
FT                   /label= OTHER
FT                   /note= "Other= Pyro-glutamic acid or 5-oxo proline"
FT   Peptide         11. .16
FT                   /note= "Spacer peptide"
FT   Peptide         17. .37
FT                   /note= "Tetanus toxoid sequence (947-967 aa)"
XX
PN   WO200185763-A2.
XX
PD   15-NOV-2001.
XX
PF   04-MAY-2001; 2001WO-US014363.
XX
PR   05-MAY-2000; 2000US-0202328P.
XX
PA   (APHT-) APHTON CORP.
XX
PI   Grimes S,  Michaeli D,  Stevens VC;
XX
DR   WPI; 2002-049440/06.
XX
```

```
PT   Novel synthetic immunogen for inducing immune response against
PT   gonadotropin releasing hormone, comprises fusion peptide having
PT   promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
PT   or its analog.
XX
PS   Claim 11; Page 9; 43pp; English.
XX
CC   The invention relates to a synthetic immunogen for inducing specific
CC   antibodies against gonadotropin releasing hormone (GnRH also known as
CC   luteinising hormone releasing hormone, LHRH) comprising a fusion peptide
CC   which comprises a promiscuous helper T-cell peptide epitope and
CC   immunomimic peptide epitope or its analogue. The synthetic immunogen is
CC   useful inducing an immune response against GnRH in an animal subject, and
CC   as such is useful as a contraceptive and in the treatment of diseases
CC   such as cancer (of the breast, uterus and other gynaecological cancer),
CC   endometriosis, uterine fibroids, benign prostatic hypertrophy and
CC   prostate cancer. The immunogen is effective in eliciting high and
CC   specific anti-GnRH antibody titres. The present sequence is a synthetic
CC   immunogen of the invention
XX
SQ   Sequence 37 AA;

  Query Match             50.0%;  Score 87;  DB 5;  Length 37;
  Best Local Similarity   100.0%;  Pred. No. 3.3e-05;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 HWSYGLRPGSSGPSL 16
              ||||||||||||||||
Db          2 HWSYGLRPGSSGPSL 16


RESULT 15
AAU11429
ID   AAU11429 standard; peptide; 50 AA.
XX
AC   AAU11429;
XX
DT   12-MAR-2002  (first entry)
XX
DE   Synthetic immunogen peptide 10.
XX
KW   Gonadotrophin releasing hormone; GnRH; synthetic immunogen;
KW   luteinising hormone releasing hormone; LHRH; contraceptive;
KW   promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
KW   breast cancer; uterine cancer; gynaecological cancer; endometriosis;
KW   uterine fibroid; benign prostatic hypertrophy; prostate cancer.
XX
OS   Clostridium tetani.
OS   Mammalia.
OS   Synthetic.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .10
FT                   /note= "Gonadotrophin releasing hormone epitope (1. .10
FT                   aa)"
FT   Misc-difference 1
FT                   /label= OTHER
FT                   /note= "Other= Pyro-glutamic acid or 5-oxo proline"
FT   Peptide         11. .16
FT                   /note= "Spacer peptide"
FT   Peptide         17. .37
FT                   /note= "Tetanus toxoid (947-967 aa)"
FT   Peptide         38. .41
FT                   /note= "Spacer peptide"
FT   Peptide         42. .50
FT                   /note= "Gonadotrophin releasing hormone epitope (2-10
FT                   aa)"
FT   Modified-site   50
FT                   /note= "Amidated glycine or glycinamide"
XX
PN   WO200185763-A2.
XX
PD   15-NOV-2001.
XX
PF   04-MAY-2001; 2001WO-US014363.
XX
PR   05-MAY-2000; 2000US-0202328P.
XX
PA   (APHT-) APHTON CORP.
XX
PI   Grimes S,  Michaeli D,  Stevens VC;
XX
DR   WPI; 2002-049440/06.
XX
PT   Novel synthetic immunogen for inducing immune response against
PT   gonadotropin releasing hormone, comprises fusion peptide having
PT   promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
PT   or its analog.
XX
PS   Claim 11; Page 11; 43pp; English.
XX
CC   The invention relates to a synthetic immunogen for inducing specific
CC   antibodies against gonadotropin releasing hormone (GnRH also known as
CC   luteinising hormone releasing hormone, LHRH) comprising a fusion peptide
CC   which comprises a promiscuous helper T-cell peptide epitope and
CC   immunomimic peptide epitope or its analogue. The synthetic immunogen is
CC   useful inducing an immune response against GnRH in an animal subject, and
CC   as such is useful as a contraceptive and in the treatment of diseases
CC   such as cancer (of the breast, uterus and other gynaecological cancer),
CC   endometriosis, uterine fibroids, benign prostatic hypertrophy and
CC   prostate cancer. The immunogen is effective in eliciting high and
CC   specific anti-GnRH antibody titres. The present sequence is a synthetic
CC   immunogen of the invention
XX
SQ   Sequence 50 AA;

  Query Match             50.0%;  Score 87;  DB 5;  Length 50;
  Best Local Similarity   100.0%;  Pred. No. 4.8e-05;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 HWSYGLRPGSSGPSL 16
              ||||||||||||||||
Db          2 HWSYGLRPGSSGPSL 16


Search completed: March 10, 2004, 09:12:11
Job time : 51.1984 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:54 ; Search time 33.6031 Seconds
(without alignments)
319.245 Million cell updates/sec

Title: US-09-848-834A-13
Perfect score: 174
Sequence: 1 XHWSYGLRPGSSGPSLKLLSEIKGVIVHRLEGVE 34

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	72	41.4	534	12	Q04243	Q04243 measles vir
2	72	41.4	537	12	Q04242	Q04242 measles vir
3	72	41.4	545	12	Q9PXA4	Q9pxa4 measles vir
4	72	41.4	546	12	Q91HA5	Q91ha5 rinderpest
5	72	41.4	550	12	P90331	P90331 measles vir
6	72	41.4	550	12	Q9QEX0	Q9qex0 measles vir
7	72	41.4	550	12	Q9QEW9	Q9qew9 measles vir
8	72	41.4	550	12	P90330	P90330 measles vir
9	72	41.4	550	12	Q9QEW7	Q9qew7 measles vir
10	72	41.4	550	12	Q9WMK4	Q9wmk4 measles vir
11	72	41.4	550	12	Q89495	Q89495 measles vir
12	72	41.4	550	12	Q8V049	Q8v049 measles vir
13	72	41.4	550	12	Q9YJ94	Q9yj94 measles vir
14	72	41.4	550	12	Q9QEX1	Q9qex1 measles vir
15	72	41.4	550	12	Q9QEW8	Q9qew8 measles vir
16	72	41.4	553	12	O93055	O93055 measles vir
17	72	41.4	553	12	Q9IC36	Q9ic36 measles vir
18	72	41.4	553	12	P88973	P88973 measles vir
19	72	41.4	553	12	Q83536	Q83536 measles vir
20	72	41.4	553	12	O11383	O11383 measles vir
21	72	41.4	553	12	Q9IFK2	Q9ifk2 measles vir
22	72	41.4	553	12	Q83533	Q83533 measles vir
23	72	41.4	553	12	Q83525	Q83525 measles vir
24	72	41.4	553	12	Q83518	Q83518 measles vir
25	72	41.4	553	12	P88974	P88974 measles vir
26	72	41.4	553	12	Q83527	Q83527 measles vir
27	72	41.4	553	12	Q83521	Q83521 measles vir
28	72	41.4	553	12	Q83530	Q83530 measles vir
29	72	41.4	553	12	O91248	O91248 measles vir
30	72	41.4	553	12	Q91QP2	Q91qp2 measles vir
31	72	41.4	553	12	Q04244	Q04244 measles vir
32	72	41.4	579	12	Q9PWU4	Q9pwu4 measles vir
33	68	39.1	545	12	Q9QEW6	Q9qew6 measles vir
34	68	39.1	553	12	O11380	O11380 measles vir
35	66	37.9	528	12	Q9YJW9	Q9yjw9 canine dist
36	66	37.9	530	12	Q8QV06	Q8qv06 canine dist
37	66	37.9	662	12	Q9DXZ2	Q9dxz2 canine dist
38	66	37.9	662	12	Q91KN3	Q91kn3 canine dist
39	66	37.9	662	12	Q9YKL7	Q9ykl7 canine dist
40	66	37.9	662	12	O89327	O89327 canine dist
41	65.5	37.6	87	13	Q9YI26	Q9yi26 sparus aura
42	65	37.4	552	12	Q66147	Q66147 cetacean mo
43	64.5	37.1	68	13	Q8JIF4	Q8jif4 acanthopagr
44	62.5	35.9	64	13	Q8JIF2	Q8jif2 pagrus majo
45	62	35.6	552	12	Q66409	Q66409 dolphin mor

ALIGNMENTS

RESULT 1
Q04243
ID Q04243 PRELIMINARY; PRT; 534 AA.
AC Q04243;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Fusion protein.
GN F.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11234;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89003063; PubMed=3167982;
RA Cattaneo R., Schmid A., Eschle D., Baczko K., ter Meulen V.,
RA Billeter M.A.;
RT "Biased hypermutation and other genetic changes in defective measles
RT viruses in human brain infections.";
RL Cell 55:255-265(1988).
RN [2]
RP SEQUENCE FROM N.A.
RA Cattaneo R., Billeter M.A.;
RL Virology 0:0-0(0).
DR EMBL; X16568; CAA34581.1; -.
DR EMBL; X16568; CAA34582.1; -.
DR HSSP; P04849; 1SVF.
DR GO; GO:0019039; F:viral-cell fusion molecule activity; IEA.
DR GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.
SQ SEQUENCE 534 AA; 57899 MW; 637245E23B5BE044 CRC64;

Query Match 41.4%; Score 72; DB 12; Length 534;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 LSEIKGVIVHRLEGV 33

```
                  ||||||||||||||||
Db        291 LSEIKGVIVHRLEGV 305


RESULT 2
Q04242
ID   Q04242      PRELIMINARY;      PRT;   537 AA.
AC   Q04242;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Fusion protein.
GN   F.
OS   Measles virus.
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC   Paramyxoviridae; Paramyxovirinae; Morbillivirus..
OX   NCBI_TaxID=11234;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89003063; PubMed=3167982;
RA   Cattaneo R., Schmid A., Eschle D., Baczko K., ter Meulen V.,
RA   Billeter M.A.;
RT   "Biased hypermutation and other genetic changes in defective measles
RT.  viruses in human brain infections.";
RL   Cell 55:255-265(1988).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Cattaneo R., Billeter M.A.;
RL   Virology 0:0-0(0).
DR   EMBL; X16567; CAA34574.1; -.
DR   EMBL; X16567; CAA34575.1; -.
DR   HSSP; P04849; 1SVF.
DR   GO; GO:0019039; F:viral-cell fusion molecule activity; IEA.
DR   GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR   InterPro; IPR000776; Fusion_gly.
DR   Pfam; PF00523; fusion_gly; 1.
SQ   SEQUENCE   537 AA;  58275 MW;  D0A60AC66D979E06 CRC64;

  Query Match             41.4%;  Score 72;  DB 12;  Length 537;
  Best Local Similarity   100.0%;  Pred. No. 0.13;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         19 LSEIKGVIVHRLEGV 33
                ||||||||||||||||
Db        291 LSEIKGVIVHRLEGV 305


RESULT 3
Q9PXA4
ID   Q9PXA4      PRELIMINARY;      PRT;   545 AA.
AC   Q9PXA4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Fusion protein.
OS   Measles virus.
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC   Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX   NCBI_TaxID=11234;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OSA-3;
RA   Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,
RA   Ogura H.;
RT   "Nucleotide sequences of the fusion protein gene of subacute
RT   sclerosing panencephalitis viruses: deduced amino acid sequences
RT   showed the cytoplasmic domain highly mutated --truncated, elongated or
RT   predicted secondary structure changed.";
RL   Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF179440; AAF02705.1; -.
DR   EMBL; AF179439; AAF02704.1; -.
DR   HSSP; P04849; 1SVF.
DR   GO; GO:0019039; F:viral-cell fusion molecule activity; IEA.
DR   GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR   InterPro; IPR000776; Fusion_gly.
DR   Pfam; PF00523; fusion_gly; 1.
SQ   SEQUENCE   545 AA;  58907 MW;  0234C28AE193E77D CRC64;

  Query Match             41.4%;  Score 72;  DB 12;  Length 545;
  Best Local Similarity   100.0%;  Pred. No. 0.14;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         19 LSEIKGVIVHRLEGV 33
                ||||||||||||||||
Db        288 LSEIKGVIVHRLEGV 302


RESULT 4
Q91HA5
ID   Q91HA5      PRELIMINARY;      PRT;   546 AA.
AC   Q91HA5;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Fusion protein.
GN   F.
OS   Rinderpest virus.
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC   Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX   NCBI_TaxID=11241;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K;
RX   MEDLINE=21014265; PubMed=11186456;
RA   Aianot P.K., Sminev A.G., Bezborodova S.V., Starov S.K., Drygin V.V.,
RA   Gusev A.A.;
RT   "Primary structure of the F-gene from Rinderpest virus strain K.";
RL   Mol. Gen. Mikrobiol. Virusol. 4:29-33(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K;
RA   Ayanot P.K., Sminev A.G., Bezborodova S.V., Starov S.K., Drygin V.V.,
RA   Gusev A.A.;
RL   Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY035887; AAK63190.1; -.
DR   PIR; PQ0866; PQ0866.
DR   PIR; PQ0867; PQ0867.
DR   PIR; PQ0873; PQ0873.
DR   GO; GO:0019039; F:viral-cell fusion molecule activity; IEA.
DR   GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR   InterPro; IPR000776; Fusion_gly.
DR   Pfam; PF00523; fusion_gly; 1.
SQ   SEQUENCE   546 AA;  58572 MW;  449B2B2DD7405F0B CRC64;

  Query Match             41.4%;  Score 72;  DB 12;  Length 546;
  Best Local Similarity   100.0%;  Pred. No. 0.14;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         19 LSEIKGVIVHRLEGV 33
                ||||||||||||||||
Db        284 LSEIKGVIVHRLEGV 298


RESULT 5
P90331
ID   P90331      PRELIMINARY;      PRT;   550 AA.
AC   P90331;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Fusion protein.
GN   F.
OS   Measles virus.
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC   Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX   NCBI_TaxID=11234;
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NAGAHATA;
RA   Sheng J., Watanabe M., Ueda S.;
RT   "Selection of a neurotropic variant of measles virus.";
RL   Submitted (AUG-1995) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NAGAHATA;
RA   Sheng J., Nakanishi M., Watanabe M., Ueda S.;
RT   "An amino acid alteration of F protein responsible for the enhanced
RT   fusogenicity of measles virus.";
RL   Submitted (AUG-1995) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NAGAHATA;
RA   Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,
RA   Ogura H.;
RT   "Nucleotide sequences of the fusion protein gene of subacute
RT   sclerosing panencephalitis viruses: deduced amino acid sequences
RT   showed the cytoplasmic domain highly mutated --truncated, elongated or
RT   predicted secondary structure changed.";
RL   Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; D63926; BAA09958.1; -.
DR   EMBL; AF179431; AAF02696.1; -.
DR   PIR; PQ0376; PQ0376.
DR   HSSP; P04849; 1SVF.
DR   GO; GO:0019039; F:viral-cell fusion molecule activity; IEA.
DR   GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR   InterPro; IPR000776; Fusion_gly.
DR   Pfam; PF00523; fusion_gly; 1.
SQ   SEQUENCE   550 AA;  59530 MW;  97C991C7E2169839 CRC64;

  Query Match             41.4%;  Score 72;  DB 12;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.14;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         19 LSEIKGVIVHRLEGV 33
              |||||||||||||||
Db        288 LSEIKGVIVHRLEGV 302


RESULT 6
Q9QEX0
ID   Q9QEX0      PRELIMINARY;      PRT;   550 AA.
AC   Q9QEX0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Fusion protein.
OS   Measles virus.
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC   Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX   NCBI_TaxID=11234;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Toyoshima;
RA   Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,
RA   Ogura H.;
RT   "Nucleotide sequences of the fusion protein gene of subacute
RT   sclerosing panencephalitis viruses: deduced amino acid sequences
RT   showed the cytoplasmic domain highly mutated --truncated, elongated or
RT   predicted secondary structure changed.";
RL   Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF179432; AAF02697.1; -.
DR   PIR; PQ0376; PQ0376.
DR   HSSP; P04849; 1SVF.
DR   GO; GO:0019039; F:viral-cell fusion molecule activity; IEA.
DR   GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR   InterPro; IPR000776; Fusion_gly.
DR   Pfam; PF00523; fusion_gly; 1.
SQ   SEQUENCE   550 AA;  59504 MW;  2AA969D37FA5CA17 CRC64;

  Query Match             41.4%;  Score 72;  DB 12;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.14;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         19 LSEIKGVIVHRLEGV 33
              |||||||||||||||
Db        288 LSEIKGVIVHRLEGV 302


RESULT 7
Q9QEW9
ID   Q9QEW9      PRELIMINARY;      PRT;   550 AA.
AC   Q9QEW9;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Fusion protein.
OS   Measles virus.
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC   Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX   NCBI_TaxID=11234;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OSA-2;
RA   Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,
RA   Ogura H.;
RT   "Nucleotide sequences of the fusion protein gene of subacute
RT   sclerosing panencephalitis viruses: deduced amino acid sequences
RT   showed the cytoplasmic domain highly mutated --truncated, elongated or
RT   predicted secondary structure changed.";
RL   Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF179436; AAF02701.1; -.
DR   PIR; PQ0376; PQ0376.
DR   HSSP; P04849; 1SVF.
DR   GO; GO:0019039; F:viral-cell fusion molecule activity; IEA.
DR   GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR   InterPro; IPR000776; Fusion_gly.
DR   Pfam; PF00523; fusion_gly; 1.
SQ   SEQUENCE   550 AA;  59405 MW;  0AE6DBFC5DD22BBA CRC64;

  Query Match             41.4%;  Score 72;  DB 12;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.14;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         19 LSEIKGVIVHRLEGV 33
              |||||||||||||||
Db        288 LSEIKGVIVHRLEGV 302


RESULT 8
P90330
ID   P90330      PRELIMINARY;      PRT;   550 AA.
AC   P90330;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Fusion protein.
GN   F.
OS   Measles virus.
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC   Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX   NCBI_TaxID=11234;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Nagahata(HB);
RA   Sheng J., Watanabe M., Ueda S.;
RT   "Selection of a neurotropic variant of measles virus.";
RL   Submitted (AUG-1995) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Nagahata(HB);
RA   Sheng J., Nakanishi M., Watanabe M., Ueda S.;
RT   "An amino acid alteration of F protein responsible for the enhanced
```

```
RT   fusogenicity of measles virus.";
RL   Submitted (AUG-1995) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; D63924; BAA09951.1; -.
DR   PIR; PQ0376; PQ0376.
DR   HSSP; P04849; 1SVF.
DR   GO; GO:0019039; F:viral-cell fusion molecule activity; IEA.
DR   GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR   InterPro; IPR000776; Fusion_gly.
DR   Pfam; PF00523; fusion_gly; 1.
SQ   SEQUENCE   550 AA;  59589 MW;  73E7BD457ABA39B7 CRC64;

  Query Match             41.4%;  Score 72;  DB 12;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.14;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         19 LSEIKGVIVHRLEGV 33
              |||||||||||||||
Db        288 LSEIKGVIVHRLEGV 302


RESULT 9
Q9QEW7
ID   Q9QEW7      PRELIMINARY;      PRT;   550 AA.
AC   Q9QEW7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Fusion protein.
OS   Measles virus.
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC   Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX   NCBI_TaxID=11234;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OSA-2;
RA   Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,
RA   Ogura H.;
RT   "Nucleotide sequences of the fusion protein gene of subacute
RT   sclerosing panencephalitis viruses: deduced amino acid sequences
RT   showed the cytoplasmic domain highly mutated --truncated, elongated or
RT   predicted secondary structure changed.";
RL   Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF179438; AAF02703.1; -.
DR   PIR; PQ0376; PQ0376.
DR   HSSP; P04849; 1SVF.
DR   GO; GO:0019039; F:viral-cell fusion molecule activity; IEA.
DR   GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR   InterPro; IPR000776; Fusion_gly.
DR   Pfam; PF00523; fusion_gly; 1.
SQ   SEQUENCE   550 AA;  59333 MW;  086E51FED5582BBA CRC64;

  Query Match             41.4%;  Score 72;  DB 12;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.14;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         19 LSEIKGVIVHRLEGV 33
              |||||||||||||||
Db        288 LSEIKGVIVHRLEGV 302


RESULT 10
Q9WMK4
ID   Q9WMK4      PRELIMINARY;      PRT;   550 AA.
AC   Q9WMK4;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Fusion protein.
GN   F.
OS   Measles virus.
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC   Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX   NCBI_TaxID=11234;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WTF;
RX   MEDLINE=99329215; PubMed=10400788;
RA   Johnston I.C., Ter Meulen V., Schneider-Schaulies J.,
RA   Schneider-Schaulies S.;
RT   "A recombinant measles vaccine virus expressing wild-type
RT   glycoproteins : consequences for viral spread and cell tropism.";
RL   J. Virol. 73:6903-6915(1999).
DR   EMBL; AJ133108; CAB38075.1; -.
DR   PIR; PQ0376; PQ0376.
DR   HSSP; P04849; 1SVF.
DR   GO; GO:0019039; F:viral-cell fusion molecule activity; IEA.
DR   GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR   InterPro; IPR000776; Fusion_gly.
DR   Pfam; PF00523; fusion_gly; 1.
SQ   SEQUENCE   550 AA;  59580 MW;  8255499968B5D862 CRC64;

  Query Match             41.4%;  Score 72;  DB 12;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.14;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         19 LSEIKGVIVHRLEGV 33
              |||||||||||||||
Db        288 LSEIKGVIVHRLEGV 302


RESULT 11
Q89495
ID   Q89495      PRELIMINARY;      PRT;   550 AA.
AC   Q89495;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Fusion protein.
GN   F.
OS   Measles virus.
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC   Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX   NCBI_TaxID=11234;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92230209; PubMed=1566568;
RA   Rota J.S., Hummel K.B., Rota P.A., Bellini W.J.;
RT   "Genetic variability of the glycoprotein genes of current wild-type
RT   measles isolates.";
RL   Virology 188:135-142(1992).
DR   EMBL; M81903; AAA46422.1; -.
DR   EMBL; M81901; AAA46421.1; -.
DR   PIR; PQ0376; PQ0376.
DR   HSSP; P04849; 1SVF.
DR   GO; GO:0019039; F:viral-cell fusion molecule activity; IEA.
DR   GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR   InterPro; IPR000776; Fusion_gly.
DR   Pfam; PF00523; fusion_gly; 1.
SQ   SEQUENCE   550 AA;  59564 MW;  A78EEC9CD6268E58 CRC64;

  Query Match             41.4%;  Score 72;  DB 12;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.14;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         19 LSEIKGVIVHRLEGV 33
              |||||||||||||||
Db        288 LSEIKGVIVHRLEGV 302


RESULT 12
Q8V049
ID   Q8V049      PRELIMINARY;      PRT;   550 AA.
AC   Q8V049;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
```

```
DE   Fusion protein.
GN   F.
OS   Measles virus.
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC   Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX   NCBI_TaxID=11234;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=G954;
RX   MEDLINE=21635526; PubMed=11773423;
RA   Waku Kouomou D., Wild T.F.;
RT   "Adaptation of wild-type measles virus to tissue culture.";
RL   J. Virol. 76:1505-1509(2002).
DR   EMBL; AY059392; AAL29688.1; -.
DR   PIR; PQ0376; PQ0376.
DR   GO; GO:0019039; F:viral-cell fusion molecule activity; IEA.
DR   GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR   InterPro; IPR000776; Fusion_gly.
DR   Pfam; PF00523; fusion_gly; 1.
SQ   SEQUENCE   550 AA;  59551 MW;  9A7A4BA99E4DA8E9 CRC64;

  Query Match             41.4%;  Score 72;  DB 12;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.14;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         19 LSEIKGVIVHRLEGV 33
              |||||||||||||||
Db        288 LSEIKGVIVHRLEGV 302


RESULT 13
Q9YJ94
ID   Q9YJ94      PRELIMINARY;      PRT;   550 AA.
AC   Q9YJ94;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Fusion protein.
OS   Measles virus.
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC   Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX   NCBI_TaxID=11234;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=9301V;
RX   MEDLINE=98440529; PubMed=9765410;
RA   Takeda M., Kato A., Kobune F., Sakata H., Li Y., Shioda T., Sakai Y.,
RA   Asakawa M., Nagai Y.;
RT   "Measles virus attenuation associated with transcriptional impediment
RT   and a few amino acid changes in the polymerase and accessory
RT   proteins.";
RL   J. Virol. 72:8690-8696(1998).
DR   EMBL; AB012949; BAA33877.1; -.
DR   EMBL; AB012948; BAA33871.1; -.
DR   PIR; PQ0376; PQ0376.
DR   HSSP; P04849; 1SVF.
DR   GO; GO:0019039; F:viral-cell fusion molecule activity; IEA.
DR   GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR   InterPro; IPR000776; Fusion_gly.
DR   Pfam; PF00523; fusion_gly; 1.
SQ   SEQUENCE   550 AA;  59512 MW;  7AA4F1D117197BF9 CRC64;

  Query Match             41.4%;  Score 72;  DB 12;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.14;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         19 LSEIKGVIVHRLEGV 33
              |||||||||||||||
Db        288 LSEIKGVIVHRLEGV 302


RESULT 14
Q9QEX1
ID   Q9QEX1      PRELIMINARY;      PRT;   550 AA.
AC   Q9QEX1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Fusion protein.
OS   Measles virus.
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC   Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX   NCBI_TaxID=11234;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Masusako;
RA   Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,
RA   Ogura H.;
RT   "Nucleotide sequences of the fusion protein gene of subacute
RT   sclerosing panencephalitis viruses: deduced amino acid sequences
RT   showed the cytoplasmic domain highly mutated --truncated, elongated or
RT   predicted secondary structure changed.";
RL   Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF179430; AAF02695.1; -.
DR   PIR; PQ0376; PQ0376.
DR   HSSP; P04849; 1SVF.
DR   GO; GO:0019039; F:viral-cell fusion molecule activity; IEA.
DR   GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR   InterPro; IPR000776; Fusion_gly.
DR   Pfam; PF00523; fusion_gly; 1.
SQ   SEQUENCE   550 AA;  59559 MW;  609EE024A7E59C54 CRC64;

  Query Match             41.4%;  Score 72;  DB 12;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.14;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         19 LSEIKGVIVHRLEGV 33
              |||||||||||||||
Db        288 LSEIKGVIVHRLEGV 302


RESULT 15
Q9QEW8
ID   Q9QEW8      PRELIMINARY;      PRT;   550 AA.
AC   Q9QEW8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Fusion protein.
OS   Measles virus.
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC   Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX   NCBI_TaxID=11234;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OSA-2;
RA   Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,
RA   Ogura H.;
RT   "Nucleotide sequences of the fusion protein gene of subacute
RT   sclerosing panencephalitis viruses: deduced amino acid sequences
RT   showed the cytoplasmic domain highly mutated --truncated, elongated or
RT   predicted secondary structure changed.";
RL   Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF179437; AAF02702.1; -.
DR   PIR; PQ0376; PQ0376.
DR   HSSP; P04849; 1SVF.
DR   GO; GO:0019039; F:viral-cell fusion molecule activity; IEA.
DR   GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR   InterPro; IPR000776; Fusion_gly.
DR   Pfam; PF00523; fusion_gly; 1.
SQ   SEQUENCE   550 AA;  59315 MW;  086E51FED235EBBA CRC64;

  Query Match             41.4%;  Score 72;  DB 12;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.14;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy         19 LSEIKGVIVHRLEGV 33
              |||||||||||||||
Db        288 LSEIKGVIVHRLEGV 302
```

Search completed: March 10, 2004, 09:25:35
Job time : 33.6031 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:54 ; Search time 10.5837 Seconds
(without alignments)
309.015 Million cell updates/sec

Title: US-09-848-834A-13
Perfect score: 174
Sequence: 1 XHWSYGLRPGSSGPSLKLLSEIKGVIVHRLEGVE 34

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	72	41.4	282	2	PQ0376	cell fusion glycop
2	72	41.4	282	2	PQ0388	cell fusion glycop
3	72	41.4	534	1	JU0274	cell fusion glycop
4	72	41.4	546	2	S47300	gene F protein - r
5	72	41.4	550	1	E48556	cell fusion glycop
6	72	41.4	553	1	VGNZMV	cell fusion glycop
7	71	40.8	546	1	VGNZRK	cell fusion glycop
8	71	40.8	546	2	S47305	gene F protein - r
9	66	37.9	542	2	JQ2223	cell fusion protei
10	66	37.9	546	1	VGNZRL	cell fusion glycop
11	66	37.9	662	1	VGNZCD	cell fusion glycop
12	66	37.9	662	2	S21382	cell fusion protei
13	65	37.4	552	2	S47034	cell fusion protei
14	65	37.4	631	1	VGNZPD	cell fusion glycop
15	65	37.4	631	1	A48346	cell fusion glycop
16	64	36.8	67	2	I78541	gonadoliberin prec
17	64	36.8	92	1	RHHUG	gonadoliberin prec
18	62	35.6	89	2	I51423	gonadoliberin prec
19	60	34.5	546	2	S55386	cell fusion protei
20	58	33.3	10	1	RHPGG	gonadoliberin - pi
21	58	33.3	10	1	RHSHG	gonadoliberin - sh
22	58	33.3	90	1	RHMSG	gonadoliberin prec
23	58	33.3	92	1	RHRTG	gonadoliberin prec
24	56.5	32.5	98	2	I50739	gonadotropin-relea
25	56	32.2	636	2	S47299	gene F protein - r
26	55	31.6	92	2	I50644	gonadoliberin I pr
27	54.5	31.3	80	2	S39779	aldehyde reductase
28	54.5	31.3	249	2	A41497	36K antigen pra -
29	54	31.0	10	1	RHAQ1	gonadoliberin I -
30	52	29.9	213	1	EPBO	phosphatidylcholin
31	51	29.3	75	2	AI3191	hypothetical prote
32	51	29.3	201	2	T47113	protocatechuate 3,
33	51	29.3	338	2	H87537	D-amino acid oxida
34	51	29.3	371	2	T49908	hypothetical prote
35	50.5	29.0	240	2	D70894	probable pra prote
36	50	28.7	44	2	S28770	hypothetical prote
37	50	28.7	133	2	H86379	protein F5A9.22 [i
38	50	28.7	317	2	H97192	oligopeptide ABC-t
39	50	28.7	392	2	A69474	myo-inositol-1-pho
40	50	28.7	719	2	T52510	hypothetical prote
41	50	28.7	1132	2	AD1809	phycobilisome core
42	49.5	28.4	342	1	QQECB8	UDP-N-acetylmurama
43	49	28.2	286	2	E82202	hypothetical prote
44	49	28.2	315	2	A88043	protein C13A10.3 [
45	49	28.2	316	2	C84635	hypothetical prote

ALIGNMENTS

RESULT 1
PQ0376
cell fusion glycoprotein - measles virus (strain TT) (fragment)
C;Species: measles virus
C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 24-Nov-1999
C;Accession: PQ0376
R;Schulz, T.F.; Hoad, J.G.; Whitby, D.; Tizard, E.J.; Dillon, M.J.; Weiss, R.A.
J. Gen. Virol. 73, 1581-1586, 1992
A;Title: A measles virus isolate from a child with Kawasaki disease: sequence comparis
A;Reference number: PQ0374; MUID:92300360; PMID:1607874
A;Accession: PQ0376
A;Molecule type: genomic RNA
A;Residues: 1-282 <SCH>
C;Genetics:
A;Gene: F
C;Superfamily: parainfluenza virus cell fusion protein
C;Keywords: glycoprotein; membrane fusion

Query Match 41.4%; Score 72; DB 2; Length 282;
Best Local Similarity 100.0%; Pred. No. 0.023;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         19 LSEIKGVIVHRLEGV 33
              |||||||||||||||
Db         20 LSEIKGVIVHRLEGV 34
```

RESULT 2
PQ0388
cell fusion glycoprotein - measles virus (strain Schwarz vaccine) (fragment)
C;Species: measles virus
C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 24-Nov-1999
C;Accession: PQ0388
R;Schulz, T.F.; Hoad, J.G.; Whitby, D.; Tizard, E.J.; Dillon, M.J.; Weiss, R.A.
J. Gen. Virol. 73, 1581-1586, 1992
A;Title: A measles virus isolate from a child with Kawasaki disease: sequence comparis
A;Reference number: PQ0374; MUID:92300360; PMID:1607874
A;Accession: PQ0388
A;Molecule type: genomic RNA
A;Residues: 1-282 <SCH>
C;Genetics:
A;Gene: F
C;Superfamily: parainfluenza virus cell fusion protein
C;Keywords: glycoprotein; membrane fusion

Query Match 41.4%; Score 72; DB 2; Length 282;
Best Local Similarity 100.0%; Pred. No. 0.023;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         19 LSEIKGVIVHRLEGV 33
              |||||||||||||||
Db         20 LSEIKGVIVHRLEGV 34
```

RESULT 3
JU0274
cell fusion glycoprotein precursor - subacute sclerosing panencephalitis virus (strain Y
N;Contains: fusion glycoprotein F1; fusion glycoprotein F2
C;Species: subacute sclerosing panencephalitis virus, SSPEV
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 16-Jun-2000
C;Accession: JU0274
R;Komase, K.; Haga, T.; Yoshikawa, Y.; Sato, T.A.; Yamanouchi, K.
Virus Genes 4, 173-181, 1990
A;Title: Molecular analysis of structural protein genes of the Yamagata-1 strain of defe
A;Reference number: JU0274; MUID:90385702; PMID:1698327
A;Accession: JU0274
A;Molecule type: mRNA
A;Residues: 1-534 <KOM>
A;Cross-references: EMBL:D10548; NID:g222256; PIDN:BAA01405.1; PID:g222257
A;Note: the authors translated the codon GTA for residue 459 as Gly and GGG for residue
C;Genetics:
A;Gene: F
C;Superfamily: parainfluenza virus cell fusion protein
C;Keywords: glycoprotein; membrane fusion; transmembrane protein
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-107/Product: cell fusion glycoprotein F2 #status predicted <FF2>
F;108-534/Product: cell fusion glycoprotein F1 #status predicted <FF1>
F;498-514/Domain: transmembrane #status predicted <TMN>
F;6,29,61,67/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 41.4%; Score 72; DB 1; Length 534;
Best Local Similarity 100.0%; Pred. No. 0.046;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 LSEIKGVIVHRLEGV 33
||||||||||||||||
Db 288 LSEIKGVIVHRLEGV 302

RESULT 4
S47300
gene F protein - rinderpest virus
C;Species: rinderpest virus
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 15-Oct-1999
C;Accession: S47300; PQ0865
R;Evans, S.A.; Baron, M.D.; Chamberlain, R.W.; Goatley, L.; Barrett, T.
submitted to the EMBL Data Library, March 1994
A;Description: The complete nucleotide sequence of the fusion protein gene of the vaccin
A;Reference number: S47299
A;Accession: S47300
A;Molecule type: DNA
A;Residues: 1-546 <EVA>
A;Cross-references: EMBL:Z31656; NID:g535406; PIDN:CAA83482.1; PID:g535407
R;Chamberlain, R.W.; Wamwayi, H.M.; Hockley, E.; Shaila, M.S.; Goatley, L.; Knowles, N.J
J. Gen. Virol. 74, 2775-2780, 1993
A;Title: Evidence for different lineages of rinderpest virus reflecting their geographic
A;Reference number: PQ0865; MUID:94103786; PMID:8277286
A;Accession: PQ0865
A;Molecule type: mRNA
A;Residues: 86-191 <CHA>
C;Genetics:
A;Gene: F
C;Superfamily: parainfluenza virus cell fusion protein
C;Keywords: glycoprotein; membrane fusion; transmembrane protein

Query Match 41.4%; Score 72; DB 2; Length 546;
Best Local Similarity 100.0%; Pred. No. 0.047;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 LSEIKGVIVHRLEGV 33
||||||||||||||||
Db 284 LSEIKGVIVHRLEGV 298

RESULT 5
E48556
cell fusion glycoprotein precursor - measles virus (strain AIK-C)
C;Species: measles virus
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 16-Jul-1999
C;Accession: E48556
R;Mori, T.; Sasaki, K.; Hashimoto, H.; Makino, S.
Virus Genes 7, 67-81, 1993
A;Title: Molecular cloning and complete nucleotide sequence of genomic RNA of the AIK-C
A;Reference number: A48556; MUID:93227570; PMID:8470368
A;Accession: E48556
A;Molecule type: genomic RNA
A;Residues: 1-550 <MOR>
A;Cross-references: GB:S58435; NID:g299460; PIDN:AAB26145.1; PID:g299465
A;Note: sequence extracted from NCBI backbone (NCBIN:129264, NCBIP:129272)
C;Genetics:
A;Gene: F
C;Superfamily: parainfluenza virus cell fusion protein
C;Keywords: glycoprotein; membrane fusion; transmembrane protein
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-107/Product: cell fusion glycoprotein F2 #status predicted <FF2>
F;108-550/Product: cell fusion glycoprotein F1 #status predicted <FF1>
F;113-138/Region: hydrophobic
F;495-514/Domain: transmembrane #status predicted <TMM>
F;6,29,61,67/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 41.4%; Score 72; DB 1; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.047;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 LSEIKGVIVHRLEGV 33
||||||||||||||||
Db 288 LSEIKGVIVHRLEGV 302

RESULT 6
VGNZMV
cell fusion glycoprotein precursor - measles virus
C;Species: measles virus
C;Date: 31-Mar-1988 #sequence_revision 31-Mar-1989 #text_change 16-Jun-2000
C;Accession: A26962; A25616; PQ0380; PQ0384
R;Buckland, R.; Gerald, C.; Barker, R.; Wild, T.F.
J. Gen. Virol. 68, 1695-1703, 1987
A;Title: Fusion glycoprotein of measles virus: nucleotide sequence of the gene and comp
A;Reference number: A92794; MUID:87224816; PMID:3585281
A;Accession: A26962
A;Molecule type: mRNA
A;Residues: 1-553 <BUC>
A;Cross-references: GB:D00090; NID:g222061; PIDN:BAA00056.1; PID:g222062
A;Experimental source: strain Halle
R;Richardson, C.; Hull, D.; Greer, P.; Hasel, K.; Berkovich, A.; Englund, G.; Bellini,
Virology 155, 508-523, 1986
A;Title: The nucleotide sequence of the mRNA encoding the fusion protein of measles vir
A;Reference number: A94350; MUID:87071668; PMID:3788062
A;Accession: A25616
A;Molecule type: mRNA
A;Residues: 4-553 <RIC>
A;Cross-references: GB:M14915; NID:g331762; PIDN:AAA46423.1; PID:g331763
A;Experimental source: strain Edmonston
R;Schulz, T.F.; Hoad, J.G.; Whitby, D.; Tizard, E.J.; Dillon, M.J.; Weiss, R.A.
J. Gen. Virol. 73, 1581-1586, 1992
A;Title: A measles virus isolate from a child with Kawasaki disease: sequence comparisc
A;Reference number: PQ0374; MUID:92300360; PMID:1607874
A;Accession: PQ0380
A;Molecule type: genomic RNA
A;Residues: 272-553 <SCH1>
A;Experimental source: isolate CL
A;Accession: PQ0384
A;Molecule type: genomic RNA
A;Residues: 272-553 <SCH2>
A;Experimental source: isolate SE
C;Genetics:
A;Gene: F
C;Superfamily: parainfluenza virus cell fusion protein

```
C;Keywords: glycoprotein; membrane fusion; transmembrane protein
F;1-25/Domain: signal sequence #status predicted <SIG>
F;26-110/Product: cell fusion glycoprotein F2 #status predicted <FF2>
F;111-553/Product: cell fusion glycoprotein F1 #status predicted <FF1>
F;501-517/Domain: transmembrane #status predicted <TMN>
F;32,64,70/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             41.4%;  Score 72;  DB 1;  Length 553;
  Best Local Similarity   100.0%;  Pred. No. 0.047;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         19 LSEIKGVIVHRLEGV 33
              |||||||||||||||
Db        291 LSEIKGVIVHRLEGV 305


RESULT 7
VGNZRK
cell fusion glycoprotein precursor - rinderpest virus (strain Kabete O)
N;Contains: fusion glycoprotein F1; fusion glycoprotein F2
C;Species: rinderpest virus
C;Date: 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change 25-Oct-1996
C;Accession: A31051
R;Hsu, D.; Yamanaka, M.; Miller, J.; Dale, B.; Grubman, M.; Yilma, T.
Virology 166, 149-153, 1988
A;Title: Cloning of the fusion gene of rinderpest virus: comparative sequence analysis w
A;Reference number: A31051; MUID:88322864; PMID:3413983
A;Accession: A31051
A;Molecule type: genomic RNA
A;Residues: 1-546 <HSU>
C;Genetics:
A;Gene: F
C;Superfamily: parainfluenza virus cell fusion protein
C;Keywords: glycoprotein; membrane fusion; transmembrane protein
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-108/Product: cell fusion glycoprotein F2 #status predicted <FF1>
F;109-546/Product: cell fusion glycoprotein F1 #status predicted <FF2>
F;109-134/Domain: transmembrane #status predicted <TN1>
F;491-513/Domain: transmembrane #status predicted <TN2>
F;25,57,63,518/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             40.8%;  Score 71;  DB 1;  Length 546;
  Best Local Similarity   93.3%;  Pred. No. 0.065;
  Matches   14;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         19 LSEIKGVIVHRLEGV 33
              ||||||||:||||||
Db        284 LSEIKGVIIHRLEGV 298


RESULT 8
S47305
gene F protein - rinderpest virus
C;Species: rinderpest virus
C;Date: 20-Oct-1994 #sequence_revision 08-Sep-1995 #text_change 20-Sep-1999
C;Accession: S47305; S47301
R;Baron, M.D.; Barrett, T.
submitted to the EMBL Data Library, March 1994
A;Description: The sequence of the N and L genes of Rinderpest virus, and the 5O and 3O
A;Reference number: S47283
A;Accession: S47305
A;Molecule type: mRNA
A;Residues: 1-546 <BAR>
A;Cross-references: EMBL:Z30697; NID:g535396; PIDN:CAA83181.1; PID:g535401; EMBL:Z30700;
C;Superfamily: parainfluenza virus cell fusion protein
C;Keywords: transmembrane protein

  Query Match             40.8%;  Score 71;  DB 2;  Length 546;
  Best Local Similarity   93.3%;  Pred. No. 0.065;
  Matches   14;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         19 LSEIKGVIVHRLEGV 33
              ||||||||:||||||
Db        284 LSEIKGVIIHRLEGV 298


RESULT 9
JQ2223
cell fusion protein F0 precursor - phocine distemper virus
N;Contains: F1 and F2 chains
C;Species: phocine distemper virus
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 24-Nov-1999
C;Accession: JQ2223
R;Visser, I.K.G.; van der Heijden, R.W.J.; van de Bildt, M.W.G.; Kenter, M.J.H.; Oerve
J. Gen. Virol. 74, 1989-1994, 1993
A;Title: Fusion protein gene nucleotide sequence similarities, shared antigenic sites
e virus entity.
A;Reference number: JQ2223; MUID:93389459; PMID:8376973
A;Accession: JQ2223
A;Molecule type: mRNA
A;Residues: 1-542 <VIS>
A;Cross-references: GB:L07075
A;Note: the authors translated the codon ATC for residue 4 as Leu
C;Comment: This fusion protein F0 is cleaved into F1 and F2 chains.
C;Genetics:
A;Gene: F
C;Superfamily: parainfluenza virus cell fusion protein
C;Keywords: glycoprotein; membrane fusion; transmembrane protein
F;1-15/Domain: signal sequence #status predicted <SIG>
F;16-542/Product: fusion protein #status predicted <MAT>
F;16-99/Product: F2 chain #status predicted <F2C>
F;105-542/Product: F1 chain #status predicted <F1C>
F;105-135/Region: hydrophobic
F;486-512/Domain: transmembrane #status predicted <TMM>
F;21,53,59,397/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             37.9%;  Score 66;  DB 2;  Length 542;
  Best Local Similarity   73.7%;  Pred. No. 0.33;
  Matches   14;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         15 SLKLLSEIKGVIVHRLEGV 33
              |   |||:|||||||||| |
Db        276 SYPTLSEVKGVIVHRLEAV 294


RESULT 10
VGNZRL
cell fusion glycoprotein precursor - rinderpest virus (strain L)
N;Contains: fusion glycoprotein F1; fusion glycoprotein F2
C;Species: rinderpest virus
C;Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 16-Jul-1999
C;Accession: A28921
R;Tsukiyama, K.; Yoshikawa, Y.; Yamanouchi, K.
Virology 164, 523-530, 1988
A;Title: Fusion glycoprotein (F) of rinderpest virus: entire nucleotide sequence of th
A;Reference number: A28921; MUID:88219541; PMID:3285575
A;Accession: A28921
A;Molecule type: mRNA
A;Residues: 1-546 <TSU>
A;Cross-references: GB:M20870; NID:g333898; PIDN:AAA47399.1; PID:g333899
C;Genetics:
A;Gene: F
C;Superfamily: parainfluenza virus cell fusion protein
C;Keywords: glycoprotein; membrane fusion; transmembrane protein
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-104/Product: cell fusion glycoprotein F2 #status predicted <FG2>
F;105-546/Product: cell fusion glycoprotein F1 #status predicted <FG1>
F;109-133/Domain: transmembrane #status predicted <TN1>
F;485-513/Domain: transmembrane #status predicted <TN2>
F;25,57,63/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             37.9%;  Score 66;  DB 1;  Length 546;
  Best Local Similarity   93.3%;  Pred. No. 0.33;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         19 LSEIKGVIVHRLEGV 33
```

```
                ||||||||||||| |
Db        284 LSEIKGVIVHRLESV 298


RESULT 11
VGNZCD
cell fusion glycoprotein precursor - canine distemper virus
N;Contains: fusion protein F1; fusion protein F2
C;Species: canine distemper virus
C;Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 16-Jul-1999
C;Accession: JS0321
R;Barrett, T.; Clarke, D.K.; Evans, S.A.; Rima, B.K.
Virus Res. 8, 373-386, 1987
A;Title: The nucleotide sequence of the gene encoding the F protein of canine distemper
A;Reference number: JS0321; MUID:88129050; PMID:3433924
A;Accession: JS0321
A;Molecule type: mRNA
A;Residues: 1-662 <BAR>
A;Cross-references: GB:M21849; NID:g323241; PIDN:AAA42878.1; PID:g323242
C;Genetics:
A;Gene: F
C;Superfamily: parainfluenza virus cell fusion protein
C;Keywords: glycoprotein; membrane fusion; transmembrane protein
F;1-135/Domain: signal sequence #status predicted <SIG>
F;136-224/Product: cell fusion glycoprotein F2 #status predicted <F2P>
F;225-662/Product: cell fusion glycoprotein F1 #status predicted <F1P>
F;606-629/Domain: transmembrane #status predicted <MEM>
F;62,141,173,179,517/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             37.9%;  Score 66;  DB 1;  Length 662;
  Best Local Similarity   73.7%;  Pred. No. 0.41;
  Matches   14;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         15 SLKLLSEIKGVIVHRLEGV 33
              |   |||:|||||||||| |
Db        396 SYPTLSEVKGVIVHRLEAV 414


RESULT 12
S21382
cell fusion protein - canine distemper virus
C;Species: canine distemper virus
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 24-Nov-1999
C;Accession: S21382
R;Wild, T.F.; Bernard, A.; Spehner, D.; Villeval, D.; Drillien, R.
submitted to the EMBL Data Library, April 1992
A;Description: Vaccination of mice against canine distemper virus induced encephalitis w
A;Reference number: S21382
A;Accession: S21382
A;Status: preliminary
A;Molecule type: genomic RNA
A;Residues: 1-662 <WIL>
A;Cross-references: EMBL:X65509; NID:g58853; PIDN:CAA46481.1; PID:g58854
C;Superfamily: parainfluenza virus cell fusion protein

  Query Match             37.9%;  Score 66;  DB 2;  Length 662;
  Best Local Similarity   73.7%;  Pred. No. 0.41;
  Matches   14;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         15 SLKLLSEIKGVIVHRLEGV 33
              |   |||:|||||||||| |
Db        396 SYPTLSEVKGVIVHRLEAV 414


RESULT 13
S47034
cell fusion protein precursor - porpoise morbillivirus
N;Alternate names: F protein
C;Species: porpoise morbillivirus
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 24-Nov-1999
C;Accession: S47034
R;Bolt, G.; Gottschalck, E.; Blixenkkrone-Moeller, M.; Wishaupt, R.G.A.; Welsh, M.J.; Ea
submitted to the EMBL Data Library, July 1994
A;Description: Nucleotide sequence comparisons of the F and M genes of cetacean morbill
A;Reference number: S47034
A;Accession: S47034
A;Molecule type: mRNA
A;Residues: 1-552 <BOL>
A;Cross-references: EMBL:X80757; NID:g520639; PIDN:CAA56731.1; PID:g520640
A;Experimental source: isolate Ulster 88
A;Note: the source is designated as Cetacean morbillivirus
C;Superfamily: parainfluenza virus cell fusion protein
F;1-25/Domain: signal sequence #status predicted <SIG>
F;26-552/Product: fusion protein #status predicted <MAT>

  Query Match             37.4%;  Score 65;  DB 2;  Length 552;
  Best Local Similarity   86.7%;  Pred. No. 0.46;
  Matches   13;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         19 LSEIKGVIVHRLEGV 33
              |||:|||||||||| |
Db        290 LSEVKGVIVHRLEAV 304


RESULT 14
VGNZPD
cell fusion glycoprotein precursor - phocine distemper virus
N;Contains: fusion protein F1; fusion protein F2
C;Species: phocine distemper virus
C;Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 25-Oct-1996
C;Accession: JQ1368
R;Koevamees, J.; Blixenkrone-Moeller, M.; Sharma, B.; Oervell, C.; Norrby, E.
J. Gen. Virol. 72, 2959-2966, 1991
A;Title: The nucleotide sequence and deduced amino acid composition of the haemagglutir
A;Reference number: JQ1368; MUID:92113538; PMID:1765768
A;Accession: JQ1368
A;Molecule type: genomic RNA
A;Residues: 1-631 <KOV>
C;Genetics:
A;Gene: F
C;Superfamily: parainfluenza virus cell fusion protein
C;Keywords: glycoprotein; membrane fusion; transmembrane protein
F;1-188/Product: cell fusion glycoprotein F2 #status predicted <FP2>
F;89-106/Domain: transmembrane #status predicted <TM1>
F;189-193/Region: cleavage processing #status predicted
F;194-631/Product: cell fusion glycoprotein F1 #status predicted <FP1>
F;194-212/Domain: transmembrane #status predicted <TM2>
F;575-595/Domain: transmembrane #status predicted <TM3>
F;110,142,148,486/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             37.4%;  Score 65;  DB 1;  Length 631;
  Best Local Similarity   68.4%;  Pred. No. 0.53;
  Matches   13;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         15 SLKLLSEIKGVIVHRLEGV 33
              |   |||:|||:|||||| |
Db        365 SYPTLSEVKGVVVHRLEAV 383


RESULT 15
A48346
cell fusion glycoprotein precursor - phocine distemper virus (strain Ulster/88)
N;Contains: fusion protein F1; fusion protein F2
C;Species: phocine distemper virus
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 25-Oct-1996
C;Accession: A48346
R;Curran, M.D.; Lu, Y.J.; Rima, B.K.
Arch. Virol. 126, 159-169, 1992
A;Title: The fusion protein gene of phocine distemper virus: nucleotide and deduced ami
A;Reference number: A48346; MUID:92398437; PMID:1524494
A;Accession: A48346
A;Molecule type: mRNA
A;Residues: 1-631 <CUR>
A;Note: sequence extracted from NCBI backbone (NCBIN:113098, NCBIP:113099)
C;Genetics:
A;Gene: F
```

C;Superfamily: parainfluenza virus cell fusion protein
C;Keywords: glycoprotein; membrane fusion; transmembrane protein
F;1-188/Product: cell fusion glycoprotein F2 #status predicted <FP2>
F;89-106/Domain: transmembrane #status predicted <TM1>
F;194-631/Product: cell fusion glycoprotein F1 #status predicted <FP1>
F;194-219/Domain: transmembrane #status predicted <TM2>
F;575-595/Domain: transmembrane #status predicted <TM3>
F;110,142/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 37.4%; Score 65; DB 1; Length 631;
Best Local Similarity 68.4%; Pred. No. 0.53;
Matches 13; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

```
Qy          15 SLKLLSEIKGVIVHRLEGV 33
               |   |||:|||:||||| |
Db         365 SYPTLSEVKGVVVHRLEAV 383
```

Search completed: March 10, 2004, 09:16:49
Job time : 11.6425 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:53 ; Search time 6.2179 Seconds
(without alignments)
284.724 Million cell updates/sec

Title: US-09-848-834A-13
Perfect score: 174
Sequence: 1 XHWSYGLRPGSSGPSLKLLSEIKGVIVHRLEGVE 34

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	72	41.4	534	1	VGLF_MEASY	P26032 measles vir
2	72	41.4	546	1	VGLF_RINDB	P41360 rinderpest
3	72	41.4	550	1	VGLF_MEASA	P35973 measles vir
4	72	41.4	550	1	VGLF_MEASE	P08300 measles vir
5	71	40.8	546	1	VGLF_RINDR	P41356 rinderpest
6	66	37.9	546	1	VGLF_RINDL	P10864 rinderpest
7	66	37.9	662	1	VGLF_CDVO	P12569 canine dist
8	65.5	37.6	95	1	GON1_SPAAU	P51919 sparus aura
9	65	37.4	546	1	VGLF_RINDK	P12574 rinderpest
10	65	37.4	631	1	VGLF_PHODV	P28886 phocine dis
11	64	36.8	67	1	GON1_MACMU	P55247 macaca mula
12	64	36.8	92	1	GON1_HUMAN	P01148 homo sapien
13	64	36.8	529	1	VGLF_MEASI	P26031 measles vir
14	62.5	35.9	95	1	GON1_PAGMA	P70074 pagrus majo
15	62	35.6	89	1	GON1_XENLA	P45656 xenopus lae
16	61.5	35.3	61	1	GON1_SHEEP	Q28588 ovis aries
17	60	34.5	92	1	GON1_TUPGB	Q95335 tupaia glis
18	58	33.3	63	1	GON1_MESAU	O09163 mesocricetu
19	58	33.3	90	1	GON1_MOUSE	P13562 mus musculu
20	58	33.3	90	1	GON1_RANCA	Q90y63 rana catesb
21	58	33.3	91	1	GON1_PIG	P49921 sus scrofa
22	58	33.3	92	1	GON1_RAT	P07490 rattus norv
23	58	33.3	99	1	GON1_DICLA	Q9ia10 dicentrarch
24	57	32.8	95	1	GON1_MORSA	O73812 morone saxa
25	56.5	32.5	94	1	GON1_HAPBU	P51918 haplochromi
26	55	31.6	92	1	GON1_CHICK	P37042 gallus gall
27	54.5	31.3	249	1	PRA_MYCLE	P41484 mycobacteri
28	54	31.0	10	1	GON1_ALLMI	P37041 alligator m
29	52	29.9	213	1	PPCT_BOVIN	P02720 bos taurus
30	51.5	29.6	74	1	GON3_ONCMY	P55246 oncorhynchu
31	51.5	29.6	90	1	GON8_RANDY	Q9iau2 rana dybows
32	51	29.3	393	1	FEZ1_RAT	P97577 rattus norv
33	50.5	29.0	240	1	PRA_MYCTU	O53426 mycobacteri
34	50.5	29.0	257	1	AP1B_HUMAN	Q8ww43 homo sapien
35	50	28.7	1131	1	APCE_ANASP	P80559 anabaena sp
36	50	28.7	2114	1	MY9B_MOUSE	Q9qy06 mus musculu
37	49.5	28.4	342	1	MURB_ECOLI	P08373 escherichia
38	49	28.2	92	1	GON1_CAVPO	O54713 cavia porce
39	49	28.2	521	1	UBP3_HUMAN	Q9y6i4 homo sapien
40	48.5	27.9	359	1	PTN7_RAT	P49445 rattus norv
41	48.5	27.9	520	1	UBP3_MOUSE	Q91w36 mus musculu
42	48	27.6	80	1	GON1_CLAGA	P33439 clarias gar
43	48	27.6	91	1	GON1_ORYLA	Q9dgc8 o progonado
44	48	27.6	336	1	UL34_EBV	P03185 epstein-bar
45	48	27.6	360	1	PTN7_HUMAN	P35236 homo sapien

ALIGNMENTS

RESULT 1
VGLF_MEASY
ID VGLF_MEASY STANDARD; PRT; 534 AA.
AC P26032;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
DE Fusion glycoprotein F1].
GN F.
OS Measles virus (strain Yamagata-1) (Subacute sclerose panencephalitis
OS virus).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11239;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90385702; PubMed=1698327;
RA Komase K., Haga T., Yoshikawa Y., Sato T.A., Yamanouchi K.;
RT "Molecular analysis of structural protein genes of the Yamagata-1
RT strain of defective subacute sclerosing panencephalitis virus. IV.
RT Nucleotide sequence of the fusion gene.";
RL Virus Genes 4:173-181(1990).
CC -!- FUNCTION: This protein directs fusion of viral and cellular
CC membranes.
CC -!- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC LINKED BY A DISULFIDE BOND.
CC -!- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC family.
CC --

CC --
DR EMBL; D10548; BAA01405.1; -.
DR HSSP; P04849; 1SVF.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.
KW Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
FT SIGNAL 1 23
FT CHAIN 24 534 FUSION GLYCOPROTEIN F0.
FT CHAIN 24 112 PROTEIN F2.
FT CHAIN 113 534 PROTEIN F1.
FT TRANSMEM 113 136 POTENTIAL.
FT DOMAIN 137 494 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 495 515 POTENTIAL.
FT DOMAIN 516 534 CYTOPLASMIC (POTENTIAL).
FT DISULFID 68 195 LINKAGE BETWEEN F2 & F1 (POTENTIAL).
FT CARBOHYD 29 29 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 61 61 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 67 67 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 534 AA; 57963 MW; F5B21757E643844D CRC64;

```
  Query Match             41.4%;  Score 72;  DB 1;  Length 534;
  Best Local Similarity   100.0%;  Pred. No. 0.015;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         19 LSEIKGVIVHRLEGV 33
              |||||||||||||||
Db        288 LSEIKGVIVHRLEGV 302


RESULT 2
VGLF_RINDB
ID   VGLF_RINDB     STANDARD;      PRT;   546 AA.
AC   P41360;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
DE   Fusion glycoprotein F1].
GN   F.
OS   Rinderpest virus (strain RBT1) (RDV).
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC   Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX   NCBI_TaxID=39007;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95088609; PubMed=7996154;
RA   Evans S.A., Baron M.D., Chamberlain R.W., Goatley L., Barrett T.;
RT   "Nucleotide sequence comparisons of the fusion protein gene from
RT   virulent and attenuated strains of rinderpest virus.";
RL   J. Gen. Virol. 75:3611-3617(1994).
CC   -!- FUNCTION: This protein directs fusion of viral and cellular
CC       membranes.
CC   -!- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC       LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC       family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z31656; CAA83482.1; -.
DR   PIR; S47300; S47300.
DR   HSSP; P04849; 1SVF.
DR   InterPro; IPR000776; Fusion_gly.
DR   Pfam; PF00523; fusion_gly; 1.
KW   Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
FT   SIGNAL        1     19
FT   CHAIN        20    546       FUSION GLYCOPROTEIN F0.
FT   CHAIN        20    108       F2 PROTEIN.
FT   CHAIN       109    546       F1 PROTEIN.
FT   DOMAIN      104    108       ARG/LYS-RICH (BASIC).
FT   TRANSMEM    109    133       POTENTIAL.
FT   TRANSMEM    484    513       POTENTIAL.
FT   DOMAIN      514    517       ARG/LYS-RICH (BASIC).
FT   DISULFID     64    191       LINKAGE BETWEEN F2 & F1 (POTENTIAL).
FT   CARBOHYD     25     25       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     57     57       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     63     63       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    518    518       O-LINKED (POTENTIAL).
SQ   SEQUENCE   546 AA;  58418 MW;  38B539B89344F401 CRC64;

  Query Match             41.4%;  Score 72;  DB 1;  Length 546;
  Best Local Similarity   100.0%;  Pred. No. 0.015;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         19 LSEIKGVIVHRLEGV 33
              |||||||||||||||
Db        284 LSEIKGVIVHRLEGV 298


RESULT 3
VGLF_MEASA
ID   VGLF_MEASA     STANDARD;      PRT;   550 AA.
AC   P35973;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
DE   Fusion glycoprotein F1].
GN   F.
OS   Measles virus (strain AIK-C) (Subacute sclerose panencephalitis
OS   virus).
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC   Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX   NCBI_TaxID=36408;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93227570; PubMed=8470368;
RA   Mori T., Sasaki K., Hashimoto H., Makino S.;
RT   "Molecular cloning and complete nucleotide sequence of genomic RNA of
RT   the AIK-C strain of attenuated measles virus.";
RL   Virus Genes 7:67-81(1993).
CC   -!- FUNCTION: This protein directs fusion of viral and cellular
CC       membranes.
CC   -!- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC       LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC       family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; S58435; AAB26145.1; -.
DR   PIR; E48556; E48556.
DR   HSSP; P04849; 1SVF.
DR   InterPro; IPR000776; Fusion_gly.
DR   Pfam; PF00523; fusion_gly; 1.
KW   Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
FT   SIGNAL        1     23
FT   CHAIN        24    550       FUSION GLYCOPROTEIN F0.
FT   CHAIN        24    112       PROTEIN F2.
FT   CHAIN       113    550       PROTEIN F1.
FT   TRANSMEM    113    136       POTENTIAL.
FT   DOMAIN      137    494       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    495    515       POTENTIAL.
FT   DOMAIN      516    550       CYTOPLASMIC (POTENTIAL).
FT   DISULFID     68    195       LINKAGE BETWEEN F2 & F1 (POTENTIAL).
FT   CARBOHYD     29     29       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     61     61       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     67     67       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   550 AA;  59540 MW;  AAC4DAB92DE0D938 CRC64;

  Query Match             41.4%;  Score 72;  DB 1;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.015;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         19 LSEIKGVIVHRLEGV 33
              |||||||||||||||
Db        288 LSEIKGVIVHRLEGV 302


RESULT 4
VGLF_MEASE
ID   VGLF_MEASE     STANDARD;      PRT;   550 AA.
AC   P08300;
```

```
DT   01-AUG-1988 (Rel. 08, Created)
DT   01-AUG-1988 (Rel. 08, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
DE   Fusion glycoprotein F1].
GN   F.
OS   Measles virus (strain Edmonston) (Subacute sclerose panencephalitis
OS   virus),
OS   Measles virus (strain Halle) (Subacute sclerose panencephalitis
OS   virus),
OS   Measles virus (strain Leningrad-16) (Subacute sclerose panencephalitis
OS   virus),
OS   Measles virus (strain Edmonston-Zagreb) (Subacute sclerose
OS   panencephalitis virus),
OS   Measles virus (strain Philadelphia-26) (Subacute sclerose
OS   panencephalitis virus), and
OS   Measles virus (strain Edmonston B) (Subacute sclerose panencephalitis
OS   virus).
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC   Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX   NCBI_TaxID=11235, 11236, 70147, 70149, 70148, 70146;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Edmonston;
RX   MEDLINE=87071668; PubMed=3788062;
RA   Richardson C.D., Hull D., Greer P., Hasel K., Berkovich A.,
RA   Englund G., Bellini W.J., Rima B., Lazzarini R.A.;
RT   "The nucleotide sequence of the mRNA encoding the fusion protein of
RT   measles virus (Edmonston strain): a comparison of fusion proteins
RT   from several different paramyxoviruses.";
RL   Virology 155:508-523(1986).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Halle;
RX   MEDLINE=87224816; PubMed=3585281;
RA   Buckland R., Gerald C., Barker R., Wild T.F.;
RT   "Fusion glycoprotein of measles virus: nucleotide sequence of the
RT   gene and comparison with other paramyxoviruses.";
RL   J. Gen. Virol. 68:1695-1703(1987).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Edmonston;
RX   MEDLINE=90085790; PubMed=2596022;
RA   Cattaneo R., Schmid A., Spielhofer P., Kaelin K., Baczko K.,
RA   Meulen V., Pardowitz J., Flanagan S., Rima B.K., Udem S.A.;
RT   "Mutated and hypermutated genes of persistent measles viruses which
RT   caused lethal human brain diseases.";
RL   Virology 173:415-425(1989).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Edmonston;
RX   MEDLINE=92263801; PubMed=1585658;
RA   Schmid A., Spielhofer P., Cattaneo R., Baczko K., Ter Meulen V.,
RA   Billeter M.A.;
RT   "Subacute sclerosing panencephalitis is typically characterized by
RT   alterations in the fusion protein cytoplasmic domain of the
RT   persisting measles virus.";
RL   Virology 188:910-915(1992).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Edmonston, Leningrad-16, and Edmonston-Zagreb;
RX   MEDLINE=94249283; PubMed=8191786;
RA   Rota J.S., Wang Z.D., Rota P.A., Bellini W.J.;
RT   "Comparison of sequences of the H, F, and N coding genes of measles
RT   virus vaccine strains.";
RL   Virus Res. 31:317-330(1994).
RN   [6]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Philadelphia-26;
RX   MEDLINE=94303181; PubMed=8030232;
RA   Hummel K.B., Vanchiere J.A., Bellini W.J.;
RT   "Restriction of fusion protein mRNA as a mechanism of measles virus
RT   persistence.";
RL   Virology 202:665-672(1994).
RN   [7]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Edmonston B;
RA   Billeter M.A.;
RL   Submitted (OCT-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: This protein directs fusion of viral and cellular
CC       membranes.
CC   -!- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC       LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC       family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M14915; AAA46423.1; -.
DR   EMBL; X05597; CAA29090.1; ALT_INIT.
DR   EMBL; K01711; AAA75498.1; ALT_INIT.
DR   EMBL; K01711; AAA75499.1; -.
DR   EMBL; U03657; AAA56647.1; ALT_INIT.
DR   EMBL; U03659; AAA56649.1; ALT_INIT.
DR   EMBL; U03670; AAA56660.1; ALT_INIT.
DR   EMBL; U08416; AAA50550.1; ALT_INIT.
DR   EMBL; Z66517; CAA91367.1; ALT_INIT.
DR   EMBL; Z66517; CAA91368.1; -.
DR   HSSP; P04849; 1SVF.
DR   InterPro; IPR000776; Fusion_gly.
DR   Pfam; PF00523; fusion_gly; 1.
KW   Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
FT   SIGNAL       1     23
FT   CHAIN       24    550       FUSION GLYCOPROTEIN F0.
FT   CHAIN       24    112       PROTEIN F2.
FT   CHAIN      113    550       PROTEIN F1.
FT   TRANSMEM   113    136       POTENTIAL.
FT   DOMAIN     137    494       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   495    515       POTENTIAL.
FT   DOMAIN     516    550       CYTOPLASMIC (POTENTIAL).
FT   DISULFID    68    195       LINKAGE BETWEEN F2 & F1 (POTENTIAL).
FT   CARBOHYD    29     29       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    61     61       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    67     67       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   550 AA;  59532 MW;  7AA4F1CA82169093 CRC64;

  Query Match             41.4%;  Score 72;  DB 1;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 0.015;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         19 LSEIKGVIVHRLEGV 33
              |||||||||||||||
Db        288 LSEIKGVIVHRLEGV 302


RESULT 5
VGLF_RINDR
ID   VGLF_RINDR     STANDARD;      PRT;   546 AA.
AC   P41356;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
DE   Fusion glycoprotein F1].
GN   F.
OS   Rinderpest virus (strain RBOK) (RDV).
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC   Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX   NCBI_TaxID=36409;
RN   [1]
```

```
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95088609; PubMed=7996154;
RA   Evans S.A., Baron M.D., Chamberlain R.W., Goatley L., Barrett T.;
RT   "Nucleotide sequence comparisons of the fusion protein gene from
RT   virulent and attenuated strains of rinderpest virus.";
RL   J. Gen. Virol. 75:3611-3617(1994).
CC   -!- FUNCTION: This protein directs fusion of viral and cellular
CC       membranes.
CC   -!- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC       LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC       family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z30700; CAA83186.1; -.
DR   EMBL; Z30697; CAA83181.1; -.
DR   PIR; S47305; S47305.
DR   HSSP; P04849; 1SVF.
DR   InterPro; IPR000776; Fusion_gly.
DR   Pfam; PF00523; fusion_gly; 1.
KW   Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
FT   SIGNAL        1     19
FT   CHAIN        20    546       FUSION GLYCOPROTEIN F0.
FT   CHAIN        20    108       F2 PROTEIN.
FT   CHAIN       109    546       F1 PROTEIN.
FT   DOMAIN      104    108       ARG/LYS-RICH (BASIC).
FT   TRANSMEM    109    133       POTENTIAL.
FT   TRANSMEM    484    513       POTENTIAL.
FT   DOMAIN      514    517       ARG/LYS-RICH (BASIC).
FT   DISULFID     64    191       LINKAGE BETWEEN F2 & F1 (POTENTIAL).
FT   CARBOHYD     25     25       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     57     57       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     63     63       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    518    518       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   546 AA;  58705 MW;  ED3DF8AFFDBECB95 CRC64;

  Query Match             40.8%;  Score 71;  DB 1;  Length 546;
  Best Local Similarity   93.3%;  Pred. No. 0.021;
  Matches   14;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         19 LSEIKGVIVHRLEGV 33
              |||||||||:||||||
Db        284 LSEIKGVIIHRLEGV 298


RESULT 6
VGLF_RINDL
ID   VGLF_RINDL     STANDARD;      PRT;   546 AA.
AC   P10864;
DT   01-JUL-1989 (Rel. 11, Created)
DT   01-JUL-1989 (Rel. 11, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
DE   Fusion glycoprotein F1].
GN   F.
OS   Rinderpest virus (strain L) (RDV).
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC   Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX   NCBI_TaxID=11243;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88219541; PubMed=3285575;
RA   Tsukiyama K., Yoshikawa Y., Yamanouchi K.;
RT   "Fusion glycoprotein (F) of rinderpest virus: entire nucleotide
RT   sequence of the F mRNA, and several features of the F protein.";
RL   Virology 164:523-530(1988).
CC   -!- FUNCTION: This protein directs fusion of viral and cellular
CC       membranes.
CC   -!- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC       LINKED BY A DISULFIDE BOND.
CC   -!- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC       family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M20870; AAA47399.1; -.
DR   PIR; A28921; VGNZRL.
DR   HSSP; P04849; 1SVF.
DR   InterPro; IPR000776; Fusion_gly.
DR   Pfam; PF00523; fusion_gly; 1.
KW   Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
FT   SIGNAL        1     19
FT   CHAIN        20    546       FUSION GLYCOPROTEIN F0.
FT   CHAIN        20    108       F2 PROTEIN.
FT   CHAIN       109    546       F1 PROTEIN.
FT   DOMAIN      104    108       ARG/LYS-RICH (BASIC).
FT   TRANSMEM    109    133       POTENTIAL.
FT   TRANSMEM    484    513       POTENTIAL.
FT   DOMAIN      514    517       ARG/LYS-RICH (BASIC).
FT   DISULFID     64    191       LINKAGE BETWEEN F2 & F1 (POTENTIAL).
FT   CARBOHYD     25     25       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     57     57       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     63     63       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   546 AA;  58911 MW;  985029418F28FFB5 CRC64;

  Query Match             37.9%;  Score 66;  DB 1;  Length 546;
  Best Local Similarity   93.3%;  Pred. No. 0.11;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         19 LSEIKGVIVHRLEGV 33
              |||||||||||||:|
Db        284 LSEIKGVIVHRLESV 298


RESULT 7
VGLF_CDVO
ID   VGLF_CDVO      STANDARD;      PRT;   662 AA.
AC   P12569; Q65991;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
DE   Fusion glycoprotein F1].
GN   F.
OS   Canine distemper virus (strain Onderstepoort) (CDV).
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC   Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX   NCBI_TaxID=11233;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88129050; PubMed=3433924;
RA   Barrett T., Clarke D.K., Evans S.A., Rima B.K.;
RT   "The nucleotide sequence of the gene encoding the F protein of canine
RT   distemper virus: a comparison of the deduced amino acid sequence with
RT   other paramyxoviruses.";
RL   Virus Res. 8:373-386(1987).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93227696; PubMed=8470428;
RA   Wild T.F., Bernard A., Spehner D., Villeval D., Drillien R.;
RT   "Vaccination of mice against canine distemper virus-induced
RT   encephalitis with vaccinia virus recombinants encoding measles or
RT   canine distemper virus antigens.";
```